

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2002, 22:13:59 ; Search time 57.04 Seconds

(Without alignments)
564.899 Million cell updates/sec

Title: US-09-578-063-47
Perfect score: 2247

Sequence: 1 MLETLSPQWIVSHRAEMWLL.....IHLMOEETNLSQRCRAVL 423

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1290	57.4	399	2	lysosomal acid lip
2	1290	57.4	399	2	lysosomal acid lip
3	1164.5	51.6	395	1	triacylglycerol 11
4	1139	50.7	397	1	triacylglycerol 11
5	1119	49.8	398	2	triacylglycerol 11
6	752.5	33.5	403	2	triacylglycerol 11
7	751	33.4	411	2	triacylglycerol 11
8	730	32.5	405	2	triacylglycerol 11
9	655.5	29.2	426	2	triacylglycerol 11
10	574.5	25.6	559	2	triacylglycerol 11
11	505	22.5	443	2	triacylglycerol 11
12	478.5	21.3	1585	2	triacylglycerol 11
13	475	21.1	467	2	triacylglycerol 11
14	445	19.8	344	2	triacylglycerol 11
15	422.5	18.8	548	2	triacylglycerol 11
16	412.5	18.4	460	2	triacylglycerol 11
17	411.5	18.3	413	2	triacylglycerol 11
18	294	13.1	431	2	triacylglycerol 11
19	285	12.7	538	2	triacylglycerol 11
20	271.5	12.1	573	2	triacylglycerol 11
21	234.5	10.4	509	2	triacylglycerol 11
22	149	6.6	62	2	triacylglycerol 11
23	143	6.4	336	2	triacylglycerol 11
24	136	5.6	987	2	triacylglycerol 11
25	113	5.0	412	2	triacylglycerol 11
26	112.5	5.0	40	2	triacylglycerol 11
27	105	4.7	430	2	triacylglycerol 11
28	103.5	4.6	798	2	triacylglycerol 11
29	101.5	4.5	929	2	triacylglycerol 11

ALIGNMENTS

30	101	4.5	436	2	JC5021
31	99.5	4.4	286	2	D75217
32	99.5	4.4	1280	2	B34087
33	99.5	4.4	460	2	G96764
34	97.5	4.3	322	2	B70957
35	97.5	4.3	456	2	E69391
36	96.5	4.3	287	1	E64053
37	96.5	4.3	1275	2	I38588
38	96	4.3	809	2	S67153
39	96	4.3	934	2	B82364
40	95.5	4.3	1259	4	GNDH1
41	95	4.2	973	2	T41201
42	94.5	4.2	561	2	G83913
43	93	4.1	1275	2	B28096
44	93	4.1	2311	1	TVCRSR
45	92.5	4.1	1275	2	S65824

RESULT 1

S41408

lysosomal acid lipase (EC 3.1.1.-) / sterol esterase (EC 3.1.1.13) precursor - human

C:Species: Homo sapiens (man)

C:Date: 13-Jan-1995 #sequence, revision 13-Jan-1995 #text, change 18-Jun-1999

C:Accession: S41408; A39315; S47187

R:Amel, D.; Merkel, M.; Eckerskorn, C.; Greten, H.

Eur. J. Biochem. 219, 905-914, 1994

A:Title: Purification, characterization and molecular cloning of human hepatic lysoso

A:Reference number: S41408; M01D:94155897

A:Accession: S41408

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-399 <AME>

A:Cross-references: EMBL:X76488; NID:9434305; P1DN:CA54026.1; PID:9434306

J. Biol. Chem. 266, 22479-22484, 1991

R:Anderson, R.A.; Sando, G.N.

A:Title: Cloning and expression of cDNA encoding human lysosomal acid lipase/choleste

A:Reference number: A39315; M01D:92042192

A:Accession: A39315

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-15, 'P', 17-399 <AND>

A:Cross-references: GB:M74775; NID:9187151; P1DN:AA59519.1; PID:9187152

R:Du, H.; Gregory, G.A.

submitted to the EMBL Data Library, April 1994

A:Description: Structural conservation of putative functional motifs between mouse an

A:Reference number: S47187

A:Accession: S47187

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-22, 'R', 24-399 <DUH>

A:Cross-references: EMBL:Z11690; NID:9506430; P1DN:CA83495.1; PID:9506431

C:Genetics:

A:Gene: GDB:LIPA

A:Cross-references: GDB:120153; OMIM:278000

A:Map position: 10q24-10q25

C:Superfamily: triacylglycerol lipase, lingual

C:Keywords: carboxylic ester hydrolase; glycoprotein

Query Match 57.4%; Score 1290; DB 2; Length 399;
Best Local Similarity 61.0%; Pred. No. 6, 1e-106;
Matches 242; Conservative 59; Mismatches 94; Indels 2; Gaps 1;

QY	15	MEMMLIVAVMQRNNSVHMPK--AVDPEAFMNSITLHOGYCEERYATEEGYI 72
DB	1	MKRPFGLVLCVLTWTHSGSGKLTAVDPETNMNVSITLISWGFESEYELVETEDGYI 60
QY	73	LSYVRIPRGLVOPKRTGSRPVLLHGLVGCASWISNLPNNSIGFTLADAGFDVWNGNS 132
DB	61	LCINRIPRGRKNSDKGPKRVYVFLQRLGLADSSNWTNLANSLGFTLADAGFDVWNGNS 120

RESULT 2
G01416
lysosomal acid lipase - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 18-Jun-1997
C:Accession: G01416
R:Du, H.
submitted to the EMBL Data Library, April 1994
A:Reference number: G06919
A:Accession: G01416
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-399 <DUX>
A:Cross-references: EMBL:U08464; NID:g505052; PIDN:AA60328.1; PID:g505053
C:Superfamily: triacylglycerol lipase, lingual
C:Keywords: glycoprotein

[illegible]

Align seg 1/1 to: HSLAL from: 1 to: 2626

283	smetAaMetSerArgAlaSerValTyrValAlaHisIstHLeuAlaGly	299
957	ATTTAATATGTCTAGAGTGAGTGTATACAAACATTCCTCGCTGA	1006
300	ThisSerValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGly	316
1007	ACTTCTGTGCAAAACATGTACACCTGGACCAGCCTGTTAAATCCAAA	1056
316	wLeuArgAlaPheAspTrpGlySerGluThrLysAsnLeuGluLysCys	333
1057	GTTTCACGCCCTTGTGACTGGGGAGAAGCACGTCCAAGATATTTTCATACA	1106
333	sngInProThrProValArgTyrArgValArgAspMetThrValProThr	349
1107	ACCAGAGTATCTCTCCACCATACAAATGTAAAGAACATCTGTCGCCACT	1156
350	AlaMetTrpThrGlyGlyGlnAspTrpLeuSerAsnProGluAspVally	366
1157	GCAGCTCGAGACGGGGGCTGCAGCTGGCTTGACAGTGTCTACAGCGCAA	1206
366	sMeLeuLeuSerGluValThrAsnLeuIleuTyrisLysAsnIleProc	383
1207	TATCTTACTGACTGATGCACCAACTGGGTGTTCATAGAGCAATTCGG	1256
383	IuTPraHisValAspPheIleTrpGlyLeuAspAlaProHisArgMet	399
1257	AATGGAGCACTTGACTTATTTGGGGCTCGAGATCCCCCTTGAGAGCTT	1306
400	TyrAsnGluIleIleHisLeuMetGlnGln	409
1307	TATAATAAATTTATTAATCAATGAGCAAA	1336
_name:	gb_pr:HDMlIPCHL	
..documentation_block:		
SUBJECT	HUMlIPCHL	2493 bp mRNA PRI 07-JAN-1995
DEFINITION	Human lysosomal acid lipase/cholesteryl esterase mRNA, complete cds.	
KEYWORDS	M74775.1 GI:187151	
ORGANISM	Lysosomal acid lipase/cholesteryl esterase. Homo sapiens fibroblast CDNA to mRNA. Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostom Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 2493) Anderson,R.A. and Sando,G.N. Cloning and expression of cDNA encoding human lysosomal acid lipase/cholesteryl ester hydrolase. Similarities to gastric and intestinal lipases J Biol Chem. 266 (33), 22479-22484 (1991)	
SDS	92042192	
location/Qualifiers		
1..2493	/organism="Homo sapiens"	
/db_xref="taxon:9606"		
/tissue_type="fibroblast"		
/map="10"		
<1..2493	/gene="LIPA"	
/note="G00-120-153"		
1..2493	/gene="LIPA"	
41..1240	/gene="LIPA"	
/EC_number="3.1.1.13"		
/note="alternate initiator ARG may be at position 47"		
/codon_start=1		
/db_xref="GDB:G00-120-153"		
/product="lysosomal acid lipase/cholesteryl esterase"		
/protein_10="AA59519.1"		
/db_xref="GI:187152"		

```

/translation="MKRFLGLVCLVLPMLHSEGSGLKFLAVDPETNNVSEIISYM
GPSESELYVEDGYILCLNRIIPHGKNSDKPRKPVFLQGLADSSNWTNLANLS
SLGFIADAGPDMWMSNRGNRSRKHRTLSYSDPFAFSYDEAKKXOLPASINIL
NKTGEOYVYVHSGGTITGFIASFQIPELARKIKRFFLGPVASYAFCTSPAKLGR
LPDHLIKDLKDEKEFLPQSAFLKWLGTHTVHLIKELGNCFLCIGPNERNLNSR
VYVYTHSPAGTSVONMLHMSQAVKFOKQAFDMGSSANFYHNSQSPPTVYKDL
VPTAVWSGGHMDLADYDVNILLTQITNLYVFHESIPEWEHLDFIMCLDAPWRLYNKI
NLMRKYQ"
41..103
/gene="LIPA"
/note="alternate signal peptide may begin at position 47;
as predicted from von Heijne, G. (1986) Nucl. Acids Res.
14,4683-4690; G00-120-153"
mat_peptide
104..1237
/gene="LIPA"
/EC_number="3.1.1.13"
/note="G00-120-153"
/product="lysosomal acid lipase/cholesterol esterase"
2466..2471
/gene="LIPA"
/note="G00-120-153"
2493
/gene="LIPA"
/note="G00-120-153"
polyA_signal
/polyA_site
/gene="LIPA"
/note="G00-120-153"
BASE COUNT 692 a 484 c 516 g 801 t
ORIGIN
alignment_scores:
Quality: 1306.00 Length: 410
Ratio: 3.841 Gaps: 2
Percent Similarity: 82.927 Percent Identity: 60.732
alignment_block:
US-09-578-063-47 x HMDLIPCHL ...
Align seg 1/1 to: HMDLIPCHL from: 1 to: 2493
2 LeuGluThrLeuSerArgGlnTrpIleValSerHisArgMetGluMetTr 18
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
8 CTCGAGACAGCGCGCGGAG.....GACAGCTCCAGATGAAGAATGCG 51
18 PleuLeuIleuValAlaIleTyrMetPheGlnArgAsnValAsnSerValH 35
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
52 GTTCTTGCGGTTGGTGTCTGTGTTGTTCTGTGCGCCCTGCATCTGAGG 101
35 IsMetProThrLys.....AlaValAspProGluAlaPheMetAsnIle 49
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
102 GGTCTGGAGGAAACTGACAGCTGTGATCCTGAACAACATGATGTG 151
50 SerGluIleIleGlnHisGlnGlyTyrProCysGluGluTyrGluValAl 66
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
152 AGTGAATATATCTCTACTGCGGATTCCTAGTGAGAAATACATGTTGA 201
66 aThrGluAspGlyTyrIleLeuSerValAsnArgIleProArgGlyLeuV 83
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
202 GACGAGAAATGATATATCTGTGCTTAACCGAATTCCTCATGTGAGAGA 251
83 aGlnProIleLysThrGlySerArgProValIleLeuLeuGlnHisGly 99
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
252 AGAACCATCTGACAAAGGTCCCAACCATGTTGTTCTTCTGACATATGGC 301
100 LeuValGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerIle 116
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
302 TTGCTGCGCAGATCTAGTAACCTGGGTCAACAACCTTGCCACAGAGCT 351
116 uGlyPheIleLeuAlaAspAlaGlyPheAspValTrpMetGlyAsnSerA 133
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
352 GGGCTTCCTTCCTGCTGATGCTGTTTGACGTGAGTGGCAACACACA 401
133 rGlyAlaAlaTrpSerArgLysHisLysThrLeuSerIleAspGlnAsp 149
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
402 GAGGAATACCTGGTCTCGAACAATAGACACTCTCAGTTTCTCAGAT 451

```

```

150 GluPheTrpAlaPheSerTyrAspGluMetAlaArgPheAspLeuProAl 166
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
452 GAATTCGGGCTTCAGTTATGATGATGAGTGGCAAAATATGACCTACAGC 501
166 aValIleAsnPheIleLeuGlnLysThrGlnGlyLysIleTyrTyrV 183
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
502 TTCATTAATCTTCATTCGATTAATAACTGGCCGACCAACATGATATATG 551
183 aGlyTyrSerGlnGlyThrThrMetGlyPheIleAlaPheSerThrMet 199
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
552 TGGGCATTCACAGGACACCTATAGATTATAGATTTTACACAGATC 601
200 ProGluLeuAlaGlnLysIleLysMetCysPheAlaLeuAlaProIleAl 216
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
602 CCGAGCGGCGCTAAAGATTAATATGTTTTTCCCTCGGCTCGTGGC 651
216 aThrValLysHisAlaLysSerProGlyThrLysPheLeuLeuProA 233
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
652 TTCGGTGGCTTCGTACTAGCCCTATAGCCCAATTAAGAGATTAACG 701
233 sPheMetIleLysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThr 249
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
702 ATCATCTCATTAAGGACTTATTTGGACACAAGATTCTTCCCAAGAGT 751
250 ArgPheLeuArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAs 266
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
752 GCGTTTGTGAAGTGGCTGGTACCCAGCTTGACATCATCATCTGAA 801
266 pGlnIleLysSerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnA 283
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
802 GGAGCTGTGTGAAATCTCTGTTTCTGTGATTTATATAGAGAA 851
283 sMetAsnMetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGly 299
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
852 ATTTAATATATCTAGAGTGCATGATATACACATCTTCCCTGCTGA 901
300 ThrSerValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyI 316
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
902 ACTCTGTGCAAACATGTACTGAGGACGAGCTGTTAATTCACAAA 951
316 uLeuArgAlaPheAspTrpGlySerGluThrLysAsnLeuGlyCysA 333
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
952 GTTCAACGCTTTCAGTGGGAGACAGTGCACAGAAATATTTTCATTA 1001
333 sGlnIleProThrProValArgTyrArgValArgAspMetThrValProThr 349
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1002 ACCAGAGTTATCTCCACATACATGTAAGAGACATCTGTGCGGACT 1051
350 AlaMetTrpThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValL 366
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1052 GCAGCTGAGACGGGGGTCACGACTGGCTTCAGATGCTACGACGTAA 1101
366 sMetLeuLeuSerGluValIleHisLeuIleTyrHisLysAsnIleProG 383
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1102 TATCTTACTGATCAGATCACCACTTGTTGTTCCATAGAGCATTCGG 1151
383 IuTrpAlaHisValAspPheIleTrpGlyLeuAspAlaProHisArgMet 399
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1152 AATGGAGACATCTGACTTTCATTTGGGCTCGATGCCCTTGGAGGCTT 1201
400 TyrAsnGluIleIleHisLeuMetGlnGln 409
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1202 TATATAAATAATTAATCTAATGAGGAA 1231
seq_name: gb-pr:BC012287
seq_documentation_block:
LOCUS BC012287 2586 bp mRNA PRI 08-AUG-2001
DEFINITION Homo sapiens, Similar to lipase A, lysosomal acid, cholesterol
esterase (Wolman disease), clone MCC:5229 IMAGE:2900168, mRNA,
complete cds.
ACCESSION BC012287

```


10

1

333 gccacacatagcctgsgcctcattctgacagatgctcggttttgacgtgtgtgatgtygaa 392

ORGANISM NON SAPRENS

QY	333	gcacaacatagccttgggcttcattcttcggcagatgctgggtttttagagctgtgagatggggaa	392
Db	330	TCGCAACGCAAGCCTGGGCTTCATCTTCGTCGATGCTGGTTTGACCTGTGGATGGGCAA	388
QY	393	cagcaggggaaacgccttgcgtctcgaataaaccaaacacactcccaatagaaccaagatgct	452
Db	390	CAGCAGAGGAAATACCTGGTGTCTGGAAACAAATACACTCTCAGATTCTCAGATGAAATT	444
QY	453	ctgggcttcagttatgagtgaatgctaggtttgaccttcctcgtcagltgataaacttlat	512
Db	450	CTGGGCTTTCAGTTATGATGATGAGATGGCAAAATATGACCTTACACACTTCCATTAATCTCAT	508
QY	513	tttcagaaaacccggccaggaagaaagctatctatgatctgctcattccacagggcaccacat	572
Db	510	TCTGATATAAACTGCGCCAGAACAAAGTGTATTATGTGGGTCATCTCTAAGGCCACCATAT	566
QY	573	gggcttatttcattctccacacatgcacagagcttgcctcagaanaatccaaaatglatcttgc	633
Db	570	AGGTTTATATACATTTTCCAGATACCCGACACTGGCTTAAAGATTTAAATGTTTTTTCG	622
QY	633	tttagcacccatagccactcgtttaagcatgcaaaaagccccggagccaatttttgttgc	692
Db	630	CCTGGGCTCTGTGGCTTCCGCTCGGCTCTCTGTACAGGCCATATGGCCAAATTAGACAGCAT	686
QY	693	gcacagatagtagtcaagagatggttttggcaaaaagaattctcgtatccagaccagat	752
Db	690	ACCAAGATCATCTCATTTAAGCACTTATTTGGAGACAAAGAAATTTTCCGCCAGATGGCTT	749
QY	753	tctcagaacaactgttatttaaccttctgtygcgaagtgattcttgatcagatttgtagtaa	812
Db	750	TTTGAAAGTGCTGGGTACCCGACGTTTGACATCATGTACTGAAAGAGCTCTGTGAAA	806
QY	813	tataatgttaactctctgggttgatccaacaacacaatatgaacatgagccgagcaagtyt	872
Db	810	TCTCTGTTTTCTTGTGTGTGATTTAATGACAGAAATTTAAATATGTCTACAGTGGATGT	866
QY	873	atagtctgcccaactctctgttgaaactcgtgtcaaatatcttacaactgagcgagc	933
Db	870	ATATACAAACATCTCTCTGCTGSAACTTCTGTCCAAAACATGTATACACTGGAGCCAGCC	922
QY	933	agtgaatctgttgaactccggcgacttgaacttggggagtgagaccaaataatctgaaaa	992
Db	930	TGTTAAATTCCAAAGTTTCAAGGCTTTGACTGGGGAAGCGTCCAAAGATTTATTTTCA	986
QY	993	atgcaatcagccaaactcctctgaagtgtaagtgtaagagatatagagatagcaggtccctacgaat	1052
Db	990	TTAACAAACAGATTAATCTCTCCACATACAAATGTAAAGACATGCTTGTCCGATGTGACGT	1046
QY	1053	gttgagacaggaagtaggaagcttgaacttcaaatccagaagaagctgnaaatgtcgtccttga	1112
Db	1050	CTGAGCGGGGGGTACGACHTGCTGTCAGATGTCTACGAGCTCAATATCTTACTGACTCA	1106
QY	1113	ggttgacaacacttcataccaataagaatlatcttgaatgggtcagctagctgattcacctg	1172
Db	1110	GATCACCAACTTGGTGTTCATGTAGACATATCCGGAATGGAGACATCTTGACTCATTTG	1166
QY	1173	gggtttggatgctcctcaccglatatgtaacaatgaaatcatcatcctgtatgcagcagagga	1222
Db	1170	GGGGCTGTGATCCCTCTTGAGAGCTTTATTAATAAATTAATTAATTAATGAGAAATATCA	1222
QY	1233	gaccaact 1241	
Db	1230	GTGAAGCT 1238	

RESULT 3

LSLAL

LOCUS

DEFINITION H.sapiens mRNA for lysosomal acid lipase.

ACCESSION X76488

VERSION X76488.1

KEYWORDS GI:434305

lysosomal acid lipase.

2626 bp

mRNA

PRI

25-FEB-1994

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

nucleic search, using sw model

January 30, 2002, 20:27:03 : Search time 90.11 Seconds
(without alignments)
3189,442 Million cell updates/sec

US-09-578-063-46

1 atgttggaaccttgcacg.....gacggtgtgagccgcatgtg 1269

IDENTITY_NUC
Gapox 10.0 , Gapext 1.0

ed: 351203 segs, 113238999 residues

number of hits satisfying chosen parameters: 702406

num DB seq length: 0
num DB seq length: 200000000

l-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

/tabase :

Issued Patents,NA: *
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCRN.COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	444.2	35.0	1137	1 US-08-227-108-2	Sequence 2, Appli
2	444.2	35.0	1137	2 US-09-073-674-2	Sequence 2, Appli
3	444.2	35.0	1140	1 US-08-227-108-4	Sequence 4, Appli
4	444.2	35.0	1140	2 US-09-073-674-4	Sequence 4, Appli
5	444.2	35.0	1146	1 US-08-227-108-6	Sequence 6, Appli
6	444.2	35.0	1146	2 US-09-073-674-6	Sequence 6, Appli
7	444.2	35.0	1528	1 US-08-227-108-1	Sequence 1, Appli
8	444.2	35.0	1528	2 US-09-073-674-1	Sequence 1, Appli
9	48.2	3.8	178	1 US-08-751-782-2	Sequence 2, Appli
10	48.2	3.8	178	2 US-08-925-171-2	Sequence 2, Appli
11	46.4	3.7	43676	3 US-09-356-952-12	Sequence 12, Appli
12	36.6	2.9	3271	2 US-08-852-806-1	Sequence 1, Appli
13	36.6	2.9	3271	3 US-09-163-669-1	Sequence 1, Appli
14	36.6	2.9	3282	1 US-08-276-852-154	Sequence 15, App
15	36.6	2.9	3282	1 US-08-276-852-169	Sequence 16, App
16	36.6	2.9	3282	1 US-08-899-575-154	Sequence 16, App
17	36.6	2.9	3282	1 US-08-899-575-154	Sequence 16, App
18	36.6	2.9	3282	1 US-08-899-575-154	Sequence 16, App
19	36.6	2.9	3282	1 US-08-899-575-154	Sequence 16, App
20	36.6	2.9	3282	5 PCT-US95-08743-154	Sequence 15, App
21	36.6	2.9	3282	5 PCT-US95-08743-154	Sequence 15, App
22	36.6	2.9	13254	1 US-08-276-852-156	Sequence 16, App
23	36.6	2.9	13254	1 US-08-276-852-170	Sequence 17, App
24	36.6	2.9	13254	1 US-08-899-575-156	Sequence 16, App
25	36.6	2.9	13254	1 US-08-899-575-170	Sequence 17, App
26	36.6	2.9	13254	1 US-08-899-575-156	Sequence 16, App
27	36.6	2.9	13254	1 US-08-899-575-170	Sequence 17, App

28	36.6	2.9	13254	5 PCT-US95-08743-156	Sequence 15, App
29	36.6	2.9	13254	5 PCT-US95-08743-170	Sequence 17, App
30	33.6	2.6	8392	1 US-08-080-255-6	Sequence 6, Appli
31	33.6	2.6	8392	1 US-08-465-713-6	Sequence 6, Appli
32	33.6	2.6	8392	5 PCT-US93-05857-6	Sequence 6, Appli
33	32.6	2.6	7218	1 US-08-232-463-14	Sequence 14, Appli
34	32	2.5	2197	1 US-08-233-005-3	Sequence 3, Appli
35	32	2.5	2197	1 US-08-428-943-3	Sequence 3, Appli
36	32	2.5	2197	1 US-09-016-649-3	Sequence 3, Appli
37	32	2.5	2197	5 PCT-US95-04858-3	Sequence 3, Appli
38	32	2.5	8342	3 US-08-545-860D-63	Sequence 63, Appli
39	32	2.5	8342	5 PCT-US94-04496-63	Sequence 63, Appli
40	31.6	2.5	3343	4 US-08-965-762-19	Sequence 19, Appli
41	31	2.4	3212	3 US-08-673-814-1	Sequence 1, Appli
42	30.4	2.4	1529	1 US-08-336-778-1	Sequence 1, Appli
43	30.2	2.4	4982	3 US-08-699-103B-1	Sequence 1, Appli
44	30.2	2.4	10079	2 US-08-476-866-20	Sequence 20, Appli
45	30	2.4	3639	2 US-08-731-524B-26	Sequence 26, Appli

ALIGNMENTS

RESULT 1
US-08-227-108-2
Sequence 2, Application US/08227108.
Patent No. 5807726

GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Pennile & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fauncel, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 66141 PENNILE
TELEX: 66141 PENNILE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1137 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1137
US-08-227-108-2

Query Match 35.0%; Score 444.2; DB 1; Length 1137;
Best Local Similarity 62.8%; Pred. No. 1.5e-134;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

OY 125 acccagaagcattcatgaatatagtgaaatcatccaacaatcaaggtatccctgtgag 184
 DB 26 ACCCTGAAGTGCACATGATATAGTCACATGATCCTACTGCGGATACCCAGCTGAGG 85
 OY 185 aatatgaagtcgaactggaatggatatactcttctgttaacagagattccctgagcc 244
 DB 86 AATATGAAGTTGTGACCGAAGCGTTATATCCTTGATGACAGAAATTCCTATAGGA 145
 OY 245 tagtgaactgaagaagaagggttccaggcctgtgtgttaactgacgaatgacctagt 304
 DB 146 GGAATAATTCAGAGATATAGCCCGAGACCTGTTCATTTTGCMAACGGGTTGCTCG 205
 OY 305 gaggtgttagcaactggaattcccaactgcccacaatagctgtggtctatcttggcag 364
 DB 206 CATCAGCCCAAACTGATGATCTCCAACTCTCCAAACAGCGCTGGCTTCACTGCGCG 265
 OY 365 atgctgtgtttagcgtgtgtgagtgaggaacagaggggaaacgcctgtgtctgaaacaca 424
 DB 266 ACGCCGGGTACGACGTGTGGTGGGAAACAGCAGGGGCAACACTGGGCCAGAGAGATC 325
 OY 425 agaaccttccataagcaagaatgagttctgggcttccagtaatgagatgagtagt 484
 DB 326 TGTACTACTCCGCCCACTCCGCTCGAATCTGGGGCTTTCAGCTTTCAGATGGCTAAAT 385
 OY 485 ttgaccttccctgagtgatataacttatttgcagaanaacggcgcaaggaagatctatc 544
 DB 386 ATGACCTTCCGCCCACTGATCTTCAATTTGAAGAAAGCGAGACGAGCACTACACT 445
 OY 545 atgtcgtattcaagagggacacacatgggcttatttgcatttccacacatgccaagagc 604
 DB 446 ACGTTGGCCATTCACAGGGCACCACTGTTTCATCCCTTTTCCACATTCACAGC 505
 OY 605 tggctcagaanaatacaatgtatttgcctttagcaccatagccaactgttaagcatgcaa 664
 DB 506 TGGCAAAACGGATCAAAACCTTCTATGCAATGATCTCCGCTTCCACCGAGATGACACG 565
 OY 665 aaagcccgaggacaaatttgtgtcgcagagatagatagatgaagagatgttggta 724
 DB 566 AAACCTGTAAACAACTCATGCTGCTCCCTTCTCTTCAAGCTTATATTTGGA 625
 OY 725 aaaaaaatttctgtatcagaacagatttccagaacatttcttcaacttctgtgccc 784
 DB 626 ACAAATATTTCTACCCACACACCTTTGATGATATTTCTGCCACCGAGATATGCTCC 685
 OY 785 agtgtattcttgatcagattttagtaataatcatcatcttctgtgtgagatcaacaca 844
 DB 686 GCGAGACGCTGATCTCTCTGACCAAGCCCTGTTATATTTGTGATTTGACACTA 745
 OY 845 acaataigaalatgagcagagcaatgtatataatgctgcccacacttctgtgaacatctg 904
 DB 746 TGAACCTTGAACATGATGCTGTGATGTATCTGTCAATATATCAGAGAGAACATCG 805
 OY 905 tgcanaatatctacacagtgagccagagcagtgatcttctgtgtaactccggagattgact 964
 DB 806 TTCAAGAGCTGCTCCACAGCTGCTGTTAATCTGCGAAGTTCCAAAGCTTTTGACT 865
 OY 965 gggggagtgagaccaaacttctgaaaaatgcaatcagcacaactcctgtaagtagagag 1024
 DB 866 GGGGAGGCCAGTTCAGAACATGATGACATATCATCAGACAGCTGCTCTCTACTCAACC 925
 OY 1025 tcaagagatagagttccctcagaatgagtgagcagagagtgatagacagcttcaaatc 1084
 DB 926 TGACAGACATGATGTGCCAATGCGAGTGTGAGACGTTGCGCAACACTTGTGCGCAC 985
 OY 1085 cagaagacgtgaaatgtgtctctctgtgagtgagacacactcattcacaataagaatattc 1144
 DB 986 CTCACGATGTGACCTTTGCTTCCAAAGCTGCCAATCTCATTTACACAGAGAAATTC 1045
 OY 1145 ctgaatggtctcaagtgatcttcatctgtggtgttggatgtctcaccgtatgtaaatg 1204
 DB 1046 CTCCTTAAATCACTTGTGACTTATCTGGCCATGATGCCCTCAACGCGTTTTCATATG 1105

OY 1205 aatcatcatctagt 1221
 DB 1106 AATGTGTTCCATGATG 1122

RESULT 2

US-09-073-674-2

Sequence 2, Application US/09073674

Patent No. 5998189

GENERAL INFORMATION:

APPLICANT: Blanchard, Claire

APPLICANT: Benicourt, Claude

APPLICANT: Junien, Jean-Louis

TITLE OF INVENTION: Recombinant Dog Gastric Lipase

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESS: Warner-Lambert Company

STREET: 2800 Plymouth Road

CITY: Ann Arbor

STATE: Michigan

COUNTRY: U.S.A.

ZIP: 48105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/073, 674

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Crissey, Todd M.

REGISTRATION NUMBER: 37,807

REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC

TELECOMMUNICATION INFORMATION:

TELEPHONE: 734 622-7530

TELEFAX: 734 622-1553

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1137 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA to mRNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1137

US-09-073-674-2

Query Match

Best Local Similarity 62.88; Pred. No. 1.5e-134;

Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

OY 125 acccagaagcattcatgaatatagtgaaatcatccaacaatcaaggtatccctgtgag 184
 DB 26 ACCCTGAAGTGCACATGATATAGTCACATGATCCTACTGCGGATACCCAGCTGAGG 85
 OY 185 aatatgaagtcgaactggaatggatatactcttctgttaacagagattccctgagcc 244
 DB 86 AATATGAAGTTGTGACCGAAGCGTTATATCCTTGATGACAGAAATTCCTATAGGA 145
 OY 245 tagtgaactgaagaagaagggttccaggcctgtgtgttaactgacgaatgacctagt 304
 DB 146 GGAATAATTCAGAGATATAGCCCGAGACCTGTTCATTTTGCACACGCGTTGCTCG 205
 OY 305 gaggtgttagcaactggaattcccaactgcccacaatagctgtggtctatcttggcag 364
 DB 206 CATCAGCCCAAACTGATGATCTCCAACTCTCCAAACAGCGCTGGCTTCACTGCGCG 265
 OY 365 atgctgtgtttagcgtgtgtgagtgaggaacagaggggaaacgcctgtgtctgaaacaca 424

37

OM of: US-09-578-063-47 to: Issued_Patents_NA:* out_format : pfs

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

```

MODE=frame-p2n.model -DEVT=zip
-O/cgn2.1/USPTO.spool/US09578063/runat.30012002.092151.26325/app-query.fasta.1.488
-DB=issued_patents_NA -OPM=fastap -SUFFIX=ini -GAPOP=12.000
-GAPEXT=4.500 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blisum62
-TRANS=humana40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=2000000000
-USUR=US09578063.ccgcn1_181 -NCPD=6 -ICPD=3 -LONGLOG -NO_XIPXY
-WAIT -THREADS=1

```

```
Database: Issued_Patents_NA:4
Database sequences: 351203
Database length: 113238999
Search time (sec): 90.07000
```

Sequence	Strid	Orig	zScore	EScore	Len	len	Document	..	129	1137
/csgn2_6/prodata/2/1na/5A_COMB	seq:US	-08-227-108-2	1141.50	2461.53	1.9e-129	1137	!			
/csgn2_6/prodata/2/1na/5B_COMB	seq:US	-09-073-674-2	1141.50	2461.53	1.9e-129	1137	!			
/csgn2_6/prodata/2/1na/5A_COMB	seq:US	-08-227-108-4	1141.50	2461.43	1.9e-129	1140	!			
/csgn2_6/prodata/2/1na/5B_COMB	seq:US	-09-073-674-4	1141.50	2461.43	1.9e-129	1140	!			
/csgn2_6/prodata/2/1na/5A_COMB	seq:US	-08-227-108-6	1141.50	2461.43	1.9e-129	1146	!			
/csgn2_6/prodata/2/1na/5B_COMB	seq:US	-09-073-674-6	1141.50	2461.43	1.9e-129	1146	!			
/csgn2_6/prodata/2/1na/5A_COMB	seq:US	-08-227-108-1	1141.50	2457.71	3.0e-129	1558	!			
/csgn2_6/prodata/2/1na/5B_COMB	seq:US	-09-073-674-1	1141.50	2457.71	3.0e-129	1558	!			
/csgn2_6/prodata/2/1na/6A_COMB	seq:US	-09-356-992-12	271.50	510.09	9.2e-21	4676	!			
/csgn2_6/prodata/2/1na/5B_COMB	seq:US	-08-751-782-2	155.00	326.10	1.6e-10	178	!			
/csgn2_6/prodata/2/1na/5A_COMB	seq:US	-08-925-171-2	155.00	326.10	1.6e-10	178	!			
/csgn2_6/prodata/2/1na/5A_COMB	seq:US	-08-484-105-21	105.00	187.71	0.0083	1676	!			
/csgn2_6/prodata/2/1na/5A_COMB	seq:US	-08-484-106-21	105.00	187.71	0.0083	1676	!			
/csgn2_6/prodata/2/1na/5B_COMB	seq:US	-08-602-359A-24	102.00	187.29	0.0088	1041	!			
/csgn2_6/prodata/2/1na/5B_COMB	seq:US	-08-223-519-1	98.50	178.16	0.0284	1167	!			
/csgn2_6/prodata/2/1na/5A_COMB	seq:US	-08-456-96-1	98.50	178.16	0.0284	1167	!			
/csgn2_6/prodata/2/1na/5B_COMB	seq:US	-09-115-446-6	92.00	153.96	0.6318	2526	!			
/csgn2_6/prodata/2/1na/6B_COMB	seq:US	-09-115-446-6	92.00	153.96	0.6318	2526	!			
/csgn2_6/prodata/2/1na/5A_COMB	seq:US	-08-471-119A-1	92.00	116.26	79.56	46899	!			
/csgn2_6/prodata/2/1na/5A_COMB	seq:US	-08-080-386-3	91.00	153.79	0.6457	2160	!			
/csgn2_6/prodata/2/1na/5B_COMB	seq:US	-08-742-185-100A-3	91.00	153.79	0.6457	2160	!			
/csgn2_6/prodata/2/1na/6A_COMB	seq:US	-08-968-542C-1	89.00	136.17	6.19	6027	!			
/csgn2_6/prodata/2/1na/5B_COMB	seq:US	-08-753-559-2	85.00	140.54	3.55	2180	!			
/csgn2_6/prodata/2/1na/5A_COMB	seq:US	-09-210-474-2	85.00	140.54	3.53	2180	!			
/csgn2_6/prodata/2/1na/5B_COMB	seq:US	-08-602-359A-30	84.50	149.29	1.15	1017	!			
/csgn2_6/prodata/2/1na/5A_COMB	seq:US	-07-707-367-1	84.50	137.05	5.53	2625	!			
/csgn2_6/prodata/2/1na/5B_COMB	seq:US	-09-386-607-1	84.00	138.47	4.61	2160	!			
/csgn2_6/prodata/2/1na/5A_COMB	seq:US	-07-882-329-4	82.50	141.46	3.14	1329	!			
/csgn2_6/prodata/2/1na/5A_COMB	seq:US	-08-183-213-4	82.50	141.46	3.14	1329	!			
/csgn2_6/prodata/2/1na/5A_COMB	seq:US	-08-446-803-4	82.50	140.29	3.65	1455	!			
/csgn2_6/prodata/2/1na/5B_COMB	seq:US	-08-861-837-4	82.50	140.29	3.65	1455	!			
/csgn2_6/prodata/2/1na/6A_COMB	seq:US	-08-600-655-4	82.50	140.29	3.65	1455	!			
/csgn2_6/prodata/2/1na/6B_COMB	seq:US	-09-170-670-9	82.50	140.29	3.65	1455	!			
/csgn2_6/prodata/2/1na/6B_COMB	seq:US	-09-170-670-14	82.50	140.28	3.65	1455	!			
/csgn2_6/prodata/2/1na/6B_COMB	seq:US	-09-193-086-9	82.50	140.29	3.65	1455	!			
/csgn2_6/prodata/2/1na/6B_COMB	seq:US	-09-193-086-13	82.50	140.29	3.65	1455	!			
/csgn2_6/prodata/2/1na/6B_COMB										

[illegible]


```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-073-674-18

Query Match 51.28; Score 1150.5; DB 2; Length 379;

```

Query Match	51.28;	Score 1150.5;	DB 2;	Length 379;
-------------	--------	---------------	-------	-------------

Matches 213; Conservative 65; Mismatches 94; Indels 3; Gaps 1;

QY 37 PTKAVDPEAFMNISEIIQHOGYPCEEYEVA TEDGYILSVNRIPRGLVQPKKTGSRPVLL 96

Db 7 PT--NPEVNMINISQMSISYWCYPSEKEYEVVTEDGYILLEVRIRIPYCKKNSGNGRQRPVFL 63

QY 97 QHGLVGASNWISNLPNNSLGFILADAGFDVWMGNSRGNAWSRKHKTLSDQDEFWAFSY 156

Db 64 QHGLLAS\$NWISNLPNNSLAFILADAGYGVWLGN\$RGTW\$SRNLYSPDSVEFWAF\$F 123

QY 157 DEMARFDLPAVINFILOKTGQEKIYYVGSQGTMGFIASFSTPELAQKIKMYFALAPIA 216'

Db 124 DEMAKYDLPATIDFIVKETGQEKLHYVGHSGTTIGFIAFSTNPKLAERIKTFYALAPVA 183

QY 217 TVKHA KSPGTFKLLLPDMMIKGLFGKKEFLYQTRFLRQLVYYLCGOVILDQICSNIMLL 276

Db 184 TVKTKSLVNLRFIPPTMEKIIFGDKIEYPHNEFDQFLATQVCSRETLNVICSNALEII 243

QY 277 GGFNTNNMNSRASVYAHTLAGTSVQNILHWSQAVNSGELRAFDWGSSETKNLEKCNQPT 336

DB 244 CGFDSANLMSRLDVYVSHNPAGTSVQNM LHTQAVKSGNFQAFNWGSPAQN VVHENQPT 303

QY 337 PVRYRVRDMTVPTAMWTGGQDWLSNPEDVKMLLSEVTNLIYHKNIPEWAHVDFIWGLDAP 396

Db 304 P P Y N V T A M N V P I A V W S G G N D W L A D P Q D V D L L P K L S N L I Y H K E I L P Y N H L D F I W A M N A P 363

QY 397 HRMYNEIHLMOEE 411

Db 364 QEVYNEIISMAKDK 378

1

US-08-227-108-3

; Patent No. 5807726

APPLICANT: Blanchard, Claire

APPLICANT: Junten, Jean-Louis

```

; NUMBER OF SEQUENCES: 21
-----

```

ADDRESSEE: Pennle & Edmonds

CITY: New York

COUNTRY: U.S.A.

COMPUTER READABLE FORM:

```

;
COMPUTER: IBM PC compatible

```

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

APPLICATION NUMBER: US/08/227,108

CLASSIFICATION: 435

NAME: Fanucci, Allan A.

REFERENCE/DOCKET NUMBER: 7620-033

TELEPHONE: 212 790-9090

TELEX: 66141 PENNIE


```

; INFORMATION FOR SEQ ID NO: 3:
;
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 379 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
;
US-08-227-108-3

```

Query Match	50.8%	Score 1141.5	DB 1	Length 379
Best Local Similarity	56.5%	Pred. No. 2.4e-112		
Matches 210	Conservative 67	Mismatches 90	Indels 5	Gaps 3

[illegible]

RESULT 6
US-09-073-674-3
Sequence 3, Application US/09073674
Patent No. 5998189
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junten, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Wainer-Lambert Company
STREET: 2800 Plymouth Road
CITY: Ann Arbor
STATE: Michigan
COUNTRY: U.S.A.
ZIP: 48105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073.674
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clissey, Todd M.
REGISTRATION NUMBER: 37,807
REFERENCE/DOCKET NUMBER: 5072-D1-66-TWC
TELECOMMUNICATION INFORMATION:

```

1 TELEPHONE: 734 622-7530
2 TELEFAX: 734 622-1553
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
10
```

Query Match	50.8%	Score 1141.5	DB 2	Length 379
Best Local Similarity	56.5%	Pred No. 2.4e-112		
Matches 210	Conservative 67	Mismatches 90	Indels 5	Gaps 3

[illegible]

RESULT 7
 US-08-227-108-5
 Sequence 5, Application US/08227108
 Patent No. 5807726
 GENERAL INFORMATION:
 APPLICANT: Blanchard, Claire
 APPLICANT: Benicourt, Claude
 APPLICANT: Junten, Jean-Louis
 TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Penne & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/227,108
 FILING DATE: 03-APR-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Faunccl, Allan A.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

nucleic search, using sw model

January 30, 2002, 20:27:03 : Search time 90.11 Seconds
(without alignments)
3189.442 Million cell updates/sec

US-09-578-063-46

1269
1 atgttggaacctgtcaag.....gacgtgtgagcgctatgt 1269

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

ed : 351203 seqs, 113238999 residues

number of hits satisfying chosen parameters: 702406

num DB seq length: 0
num DB seq length: 2000000000

Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued_Patents_NA:*
2: /cgn2_6/prodata/2/lna/5A_COMB.seq:*
3: /cgn2_6/prodata/2/lna/5B_COMB.seq:*
4: /cgn2_6/prodata/2/lna/5A_COMB.seq:*
5: /cgn2_6/prodata/2/lna/5B_COMB.seq:*
6: /cgn2_6/prodata/2/lna/5A_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	444.2	35.0	1137	1	US-08-227-108-2
2	444.2	35.0	1137	2	US-09-073-674-2
3	444.2	35.0	1140	1	US-08-227-108-4
4	444.2	35.0	1140	2	US-09-073-674-4
5	444.2	35.0	1146	1	US-08-227-108-6
6	444.2	35.0	1146	2	US-09-073-674-6
7	444.2	35.0	1528	1	US-08-227-108-1
8	444.2	35.0	1528	2	US-09-073-674-1
9	48.2	3.8	178	1	US-08-276-852-1
10	48.2	3.8	178	2	US-09-073-674-2
11	46.4	3.7	43676	3	US-08-925-171-2
12	36.6	2.9	3271	2	US-08-852-806-1
13	36.6	2.9	3271	3	US-09-163-669-1
14	36.6	2.9	3282	1	US-08-276-852-154
15	36.6	2.9	3282	1	US-08-276-852-154
16	36.6	2.9	3282	1	US-08-899-575-154
17	36.6	2.9	3282	1	US-08-899-575-154
18	36.6	2.9	3282	1	US-08-899-575-154
19	36.6	2.9	3282	1	US-08-899-575-154
20	36.6	2.9	3282	5	PCT-US95-08743-154
21	36.6	2.9	3282	5	PCT-US95-08743-154
22	36.6	2.9	13254	1	US-08-276-852-156
23	36.6	2.9	13254	1	US-08-276-852-156
24	36.6	2.9	13254	1	US-08-899-575-156
25	36.6	2.9	13254	1	US-08-899-575-156
26	36.6	2.9	13254	1	US-08-899-575-156
27	36.6	2.9	13254	1	US-08-899-575-156

28	36.6	2.9	13254	5	PCT-US95-08743-156	Sequence 156, App
29	36.6	2.9	13254	5	PCT-US95-08743-170	Sequence 170, App
30	33.6	2.6	8392	1	US-08-080-255-6	Sequence 6, App1
31	33.6	2.6	8392	3	US-08-465-713-6	Sequence 6, App1
32	33.6	2.6	8392	5	PCT-US93-05857-6	Sequence 6, App1
33	32.6	2.6	7218	1	US-08-232-463-14	Sequence 14, App1
34	32.6	2.5	2197	3	US-08-428-943-3	Sequence 3, App1
35	32.6	2.5	2197	3	US-08-428-943-3	Sequence 3, App1
36	32.6	2.5	2197	3	US-09-016-649-3	Sequence 3, App1
37	32.6	2.5	2197	5	PCT-US95-04858-3	Sequence 3, App1
38	32.6	2.5	8342	3	US-08-545-8600-63	Sequence 63, App1
39	32.6	2.5	8342	5	PCT-US94-04496-63	Sequence 19, App1
40	31.6	2.5	3343	4	US-08-965-762-19	Sequence 1, App1
41	31.6	2.4	3212	3	US-08-673-814-1	Sequence 1, App1
42	30.2	2.4	1529	3	US-08-336-778-1	Sequence 1, App1
43	30.2	2.4	4982	3	US-08-699-1038-1	Sequence 1, App1
44	30.2	2.4	10079	2	US-08-476-866-20	Sequence 20, App1
45	30.2	2.4	3639	2	US-08-737-524B-26	Sequence 26, App1

ALIGNMENTS

RESULT 1

US-08-227-108-2
Sequence 2, Application US/08227108

Patent No. 5807726

GENERAL INFORMATION:

APPLICANT: Blanchard, Claire

APPLICANT: Benicourt, Claude

APPLICANT: Junien, Jean-Louis

TITLE OF INVENTION: Recombinant Dog Gastric Lipase

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Penne & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.75

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/227,108

FILING DATE: 03-APR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Farnucci, Allan A.

REGISTRATION NUMBER: 30,256

REFERENCE/DOCKET NUMBER: 7620-033

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 790-9090

TELEFAX: 212 869-8864/9741

TELEX: 66141 PENNE

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1137 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1137

US-08-227-108-2

Query Match 35.0%; Score 444.2; DB 1; length 1137;
Best Local Similarity 62.8%; Pred. No. 1.5e-134;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

125 accagaagcatcatgaataatagtaataatcaatcaatcaatccctgtgagg 184
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db ACCCTGAAGGTACCATGATATTAAGTCAGATGATACCTACTGGGATACCCAGCTGAGG 85
185 aatatgaagtcgaactgaagatgggtatcccttcgttaacaggaattccctgagcc 244
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db AATATGAAGTGTGACCGACACAGCGTTATATCCTTGGGATGACAGCAATTCCTTATGGGA 145
245 taatgaactaagaagaaggttccaggcctgtgtgttactgtcagagatgacctagtgtg 304
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db GGAATAATTCAGAGATATAGCCGGAGACCTGTGCAATTTTGGCAACACGGTTTGCTCG 205
305 gaggtgtcagaactgtgattccaaacctgcccacaataatagctgtggttcattcgtcagag 364
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db CATAGCCACAACTGGATCTCCAACTGCCCCAACACAGCTGCGCTTCACTCCGCGCG 265
365 atgtgtgttttgaagctgtggtatggggaacagaaggggaagccctgtgtcgaagaacaca 424
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db ACCGCGGGTACGAGCTGTGGCTGGGGAACACAGGGGCAACACCTGGGCGCAGGAAGATC 325
425 agacaactccatagaccagaatgagttcgtgttcagttatgatatgagatgtgctagt 484
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db TGTACTT.TCGCCCGACTCCGTCGAATCTGGGCTT TCAAGCTTGACAGATGGCTAAT 385
485 ttgaccttcctgcagtgatlaaacttatttgcagaagaagggccaggaagaatctatt 544
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db ATGACCTTCCGCCACCACTGATCTCACTTGAAGAAAAGGAGACAGCAAGCTACACT 445
545 atgtcgtatctacagagagccacacatgggtcttattgtatcttccacatccagagag 604
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db ACCTGGGCAATTCACAGGGGACACACATGGTTTCATGCGCTTTCACCAATCCCAAGC 505
605 tgcgtcagaataatcaaatgtatttgcctttagaaccctagccactgttaagaatgca 664
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db TGGGGAACCGATCAAACTTCTATGATAGCTCCCGTTGCCACCGTGAAGTACACG 565
665 aagaacccggagcaaatcttctgtgtcgcagatatagatcaagaagatgtgttgcga 724
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db AAACCTGTTAAACAACTCATGCTCGCTCCCTGTTCTCTTCAAGCTTATATTGGAA 625
725 aaaaagaattctgtatcagaaccagatttccagacaactgttatttacttctgtgccc 784
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db ACAAAATATTCTACCAACACACTTCTTGTATCAATTTCTGCCACGAGATATGCTCC 685
785 aggtgtattctgtatcagattgtatgaatataatcatttacttctgtgtgattcaacca 844
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db GCGAGAGGTGATCTCTCTGACGAAACGCCCTGTTTATCATTTTGGGATTTGACACTA 745
845 acaatataagacatgagccagagcaagtgtatatactgtcccaactctgtgtgaacatctg 904
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db TGAACCTTGAAACATAGTCGCTTGGATGTATCTGTCAATATCCACAGCAACATCGG 805
905 tgcacaatctctactgtgagccagagcaagtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 964
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db TTTCAGACGTGTCTCACTGTGCTCCAGGCTGTAAAGTGTGGAAATTTCCAAAGCTTTTGCT 865
965 gggggaggtgagcaaaaatctgtgaaaaatgtcaatcagccaactcctgtlaaggtacagag 1024
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db GGGGAAACCCAGTTTCAAGAACTGATGCACTATCATGAGCATGCCCTCCCTACACAAACC 925
1025 tcaagagatagagctgtcccttaacagcaatgtgagacagaggtcagagactgtcttcaatc 1084
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db TGACAGCATGATGTCCCAATGCGAGTGTGAAAGGTTGGACAGCACTTGTGCGCAGCC 985
1085 cagaagacgtgaaatgtctctcctgtgagtgacaacactcatctacataagaataatc 1144
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db CTCACGATGTGACCTTTTGCTTCCAAAGCTCCCAATTCATTTACACAGGAAGATTC 1045
1145 ctgaatgtgtcagctgaggtattcatctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1204
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db CTCCTTCAATCACTTGTGACTTTATCTGGGCAATGATGCGCCCTCAAGCGGTTTACAAATG 1105

Qy 1205 aatcatccatctgtatg 1221
1111 1111
Db 1106 AATGTGTTCCATGATG 1122

RESULT 2

US-09-073-674-2

Sequence 2, Application US/09073674

Patent No. 5998189

GENERAL INFORMATION:

APPLICANT: Blanchard, Claire

APPLICANT: Benicourt, Claude

APPLICANT: Julien, Jean-Louis

TITLE OF INVENTION: Recombinant Dog Gastric Lipase

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Warner-Lambert Company

STREET: 2800 Plymouth Road

CITY: Ann Arbor

STATE: Michigan

COUNTRY: U.S.A.

ZIP: 48105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/073.674

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Crissey, Todd M.

REGISTRATION NUMBER: 37,807

REFERENCE/DOCKET NUMBER: 5072-DI-66-TMC

TELECOMMUNICATION INFORMATION:

TELEPHONE: 734 622-7530

TELEFAX: 734 622-1553

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1137 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1137

US-09-073-674-2

Query Match 35.08; Score 444.2; DB 2; Length 1137;

Best Local Similarity 62.88; Pred. No. 1.5e-134;

Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

Qy 125 accagaagcatcatgaataatagtaataatcaatcaatcaatccctgtgagg 184
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 26 ACCCTGAAGGTACCATGATATTAAGTCAGATGATACCTACTGGGATACCCAGCTGAGG 85
185 aatatgaagtcgaactgaagatgggtatcccttcgttaacaggaattccctgagcc 244
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db AATATGAAGTGTGACCGACACAGCGTTATATCCTTGGGATGACAGCAATTCCTTATGGGA 145
245 taatgaactaagaagaaggttccaggcctgtgtgttactgtcagagatgacctagtgtg 304
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db GGAATAATTCAGAGATATAGCCGGAGACCTGTGCAATTTTGGCAACACGGTTTGCTCG 205
305 gaggtgtcagaactgtgattccaaacctgcccacaataatagctgtggttcattcgtcagag 364
1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db CATAGCCACAACTGGATCTCCAACTGCCCCAACACAGCTGCGCTTCACTCCGCGCG 265
265 atgtgtgttttgaagctgtggtatggggaacagaaggggaagccctgtgtcgaagaacaca 424

OM of: US-09-578-063-47 to: Issued_Patents_NA: out_format: pfs
Date: Jan 31, 2002 12:05 AM

About: Results were produced by the GenCore software, version 4.5.
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL-framer.p2n.model -DEV-x1p
-O/cgn2_1/USPTO.spool/US09578063/runatc_30012002_092151_26325/app-query.fasta.1.488
-DB-Issued_Patents_NA -OFMT-fasta -SUFFIX-tri -GAPOP-12.000
-GAPEXT-4.000 -MINMATCH-0.100 -LOOPCL-0.000 -LOOPEXT-0.000
-GAPOP-4.500 -CGAPEXT-0.050 -XGAPOP-10.000 -XGAPEXT-0.500
-GAPOP-6.000 -FGAPEXT-7.000 -YGAPOP-10.000 -YGAPEXT-0.500
-DEPOP-6.000 -DELEXT-7.000 -START-1 -MATRIX-Diosm62
-TRANS-human40.cdt -LIST-45 -DOCALIGN-200 -THR_SCORE-pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE-LOCAL -OUTFRMT-pfs
-NORM-ext -MINLEN=0 -MAXLEN=200000000
-USER=US09578063 @CGN1_1_81 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPHY
-WAIT -THREADS=1

Search information block:

Query: US-09-578-063-47
Query length: 423
Database: Issued_Patents_NA:
Database sequences: 351203
Database length: 113228999
Search time (sec): 90.070000

Score list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-227-108-2			1141.50	2461.53	1.9e-129
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-09-073-674-2			1141.50	2461.53	1.9e-129
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-227-108-4			1141.50	2461.49	1.9e-129
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-09-073-674-4			1141.50	2461.49	1.9e-129
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-227-108-6			1141.50	2461.43	1.9e-129
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-09-073-674-6			1141.50	2461.43	1.9e-129
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-227-108-1			1141.50	2457.71	3.0e-129
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-09-073-674-1			1141.50	2457.71	3.0e-129
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-09-356-952-12			271.50	310.09	9.2e-21
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-751-782-2			155.00	326.10	1.6e-10
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-925-171-2			155.00	187.71	0.0083
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-484-105-21			105.00	187.71	0.0083
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-484-106-21			102.00	187.29	0.0088
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-602-359A-24			98.50	178.16	0.0284
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-232-519-1			98.50	178.16	0.0284
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-456-956-1			92.00	153.99	0.6298
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-09-115-446-1			92.00	153.96	0.6318
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-09-115-446-5			92.00	116.26	79.56
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-471-119A-1			91.00	153.79	0.6457
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-080-386-3			91.00	153.79	0.6457
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-742-185-102			89.00	146.17	6.19
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-968-542C-1			85.00	140.54	3.53
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-755-559-2			85.00	140.54	3.53
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-210-474-2			84.50	149.29	1.15
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-602-359A-30			84.50	137.05	5.53
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-759-945-1			84.50	133.77	8.41
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-07-707-367-1			84.00	138.47	4.61
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-09-386-607-1			82.50	141.46	3.14
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-07-882-329-4			82.50	141.46	3.14
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-183-213-4			82.50	140.29	3.65
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-446-803-4			82.50	140.29	3.65
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-08-861-837-4			82.50	140.29	3.65
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-600-656-4			82.50	140.29	3.65
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-09-170-670-9			82.50	140.29	3.65
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-09-170-670-14			82.50	140.29	3.65
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-09-193-068-9			82.50	140.29	3.65
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-09-193-068-13			82.50	140.29	3.65
/cgn2_6/ptodata/2/ina/5B.COMB.seq:US-09-183-412-9			82.50	140.29	3.65
/cgn2_6/ptodata/2/ina/5A.COMB.seq:US-09-183-412-13			82.50	140.29	3.65

seq name: /cgn2_6/ptodata/2/ina/5A.COMB.seq:US-08-227-108-2

seq documentation block:

Sequence 2, Application US/08227108
Patent No. 5807726

GENERAL INFORMATION:

APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penille & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fannucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFO, IATION:
TELEPHONE: 212 790-9690
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNITE

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1137 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA

FEATURE:
NAME/KEY: CDS
LOCATION: 1..1137
US-08-227-108-2

alignment scores:
Quality: 1141.50
Ratio: 3.624
Percent Similarity: 84.677
Percent Identity: 56.452

US-09-578-063-47 x US-08-227-108-2
Align seg 1/1 to: US-08-227-108-2 from: 1 to: 1137

	Length:	Gaps:
37 Prothrlvslalvalasproctulalaphmetasnlleserclutlell 53	372	3
19 CCCACA.....AACCTTAAAGTGAATATTAAGTCAAGTAT 59		
53 eglntlsglntlytyrprocygsluglntlyrgluvalatathgluasp 70		
60 CACCTACTGGGATACCGACCTGAGCAATATGAAGTTCACCAAGAG 109		
70 lYTrllleuSerValaAsnArgllleProArgllYleuValcInProlys 86		

APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Julien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Warner-Lambert Company
STREET: 2800 Plymouth Road
CITY: Ann Arbor
STATE: Michigan
COUNTRY: U.S.A.
ZIP: 48105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0. Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,674
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Crissey, Todd M.
REGISTRATION NUMBER: 37,807
REFERENCE/DOCKET NUMBER: 5072-D1-66-TWC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 734 622-7530
TELEFAX: 734 622-1553
TELEX:
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acids
TYPE: amino acid

[illegible]

RESULT 5
US-08-227-108-3
Sequence 3, Application US/08227108
Patent No. 5807726
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junten, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-Apr-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Faunccl, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE

~~TELECOMMUNICATION INFORMATION:~~

 SOURCE/AGENT INFORMATION:
 NAME: Fannucci, Allan A.

alignment_block:
us-09-578-063-47 x HSLAL ..

Align seg 1/1 to: HSLAL from: 1 to: 2626

```

2 LeuGIuThrLeuSerArgIntPILeValSerHisArgMetGluMetTr 18
113 CTGAGACAGCGGCGCCGCGAG.....GACAGCTCCGAGAAATGAAAAAGCG 156
18 PLeuLeuIleLeuValAlaTyrMetPheGlnArgAsnValAsnSerValH 35
157 GTTCTTGCGGTGGTGGTCTGTTGGTCTCTGACCTCGCATTTCTAGG 206
35 IsMetProThrLys.....AlaValAspProGluAlaPheMetAsnIle 49
207 GGTCTGGAGGAAAGTACAGACAGCTGGATCTGTAACAAACATGAATGTG 256
50 SerGluIleIleGlnHisGlnGlyTyrProCysGluGluTyrGluValAl 66
257 AGTGAATTTATCTTCTACTGGGATTCCTAGTGAGGAATACCTAGTTGA 306
66 aThrGluAspGlyTyrIleLeuSerValAsnArgIleProArgGlyLeuV 83
307 GACAGAAATGATGATATTTCTGCGCTTAACCGAATTCCTCATGGAGGA 356
83 aGlnProLysLysThrGlySerArgProValValLeuLeuGlnHisGly 99
357 AGAACCATCTTGACAAAGGTCCCAACACAGTGTCTCTGCAACATGGC 406
100 LeuValGlyIleAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerIe 116
407 TTGCTGGCAATCTTCTACTAGCTGGTCACAAACCTTGCCACACACACCT 456
116 uGlyPheIleLeuAlaAspAlaGlyPheAspValTrpMetGlyAsnSera 133
457 GGGCTTATCTCTGCTCATGCTGTTTGGACGTGCGATGGCAACAGCA 506
133 rGgLYaSnAlaATrPserArgLysHisLysThrLeuSerIleAspGluAsp 149
507 GAGGAATFACCTGGTCTCGAACAATAGACACACTCTCAGTTTCTCAGAGAT 556
150 GluPheTrpAlaPheSerTyrAspGluMetAlaArgPheAspLeuProAl 166
557 GAATTCGGGCTTCACTATGATGAGATGCAAAATGATGACCTACACAGC 606
166 aValIleAsnPheIleLeuGlnLysThrGlyGlnGlyLysIleTyrTyr 183
607 TTCCATTAATCTCATCTGATAATAAACTGGCCAGAAACAAGTGTATATG 656
183 aGlyTyrSerGlnGlyThrThrMetGlyPheIleAlaPheSerThrMet 199
657 TGGGTATCTTCAGAGCACACATAGTTTATAGCATTTTCCACAGATC 706
200 ProGluLeuAlaGlnLysIleLysMetTyrPheAlaLeuAlaProIleAl 216
707 CCTGAGGTGCTAAAGATTAATGTTTTCCTGGGTCTCTGGC 756
216 aThrValLysHisAlaLysSerProGlyTyrLysPheLeuLeuLeuProA 233
757 TTCGCTGGCCTCTGTACTAGCCCTATGGCCAAATTAAGACGATTAACAG 806
233 sPMeTMeIleLysGlyLeuPheGlyLysGlyGluPheLeuTyrGlnThr 249
807 ATCATCTCATTAAGACCTATTATGGAGACAAGAATTTCTCCCCAGAGT 856
250 ArgPheLeuArgGlnLeuValIleTyrLeuGlyGlnValIleLeuAs 266
857 GCGTTTGTGAGAGTGGTGGTACCCAGTTTGCACTCATGTCATACATGAA 906
266 pGlnIleCysSerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnA 283
907 GGACCTCTGAGAAATCTCTGTTTCTTCTGTGTGATTTATATGAGAGAA 956

```

```

283 sMeTasMetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGly 299
957 ATTAATAATATGTCTAGAGTGAATGATATATACACACTTCTCTGCTGGA 1006
300 ThrSerValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGly 316
1007 ACTTCTGTGAAANACATGTATACCTGAGCCAGGCTTTAAATTCCAAAA 1056
316 uLeuArgAlaPheAspTrpGlySerGluThrLysAsnLeuGluLysCysA 333
1057 GTTCAAGCCTTTGACATGGGGAAGCAGACTGCCAGAATATTATTCATACA 1106
333 sncGlnProThrProValArgTyrArgValArgAspMetThrValProThr 349
1107 ACCAGAGTTATCTCCACATACAAATGTGAAGCAAGCTTGCCGCACT 1156
350 AlMetThrThrGlyGlyGlnAspTrpLeuSerAsnProGluAspVally 366
1157 GCACTCTGGAGCGGGGTCAAGACTGCTGCGATGTCTACAGCTCA 1206
366 sMeTleuLeuSerGluValThrAsnLeuIleTyrHisLysAsnIleProG 383
1207 TACTTATCTGACTCAGATCAACAACTGATGTTCCATGAGACATTCGCG 1256
383 lUtrPAlaHisValAspPheIleTrpGlyLeuAspAlaProHisArgMet 399
1257 AATGGAGCATCTTGACTTATTTGGGCGCTGATGCCCTTGCGAGGCTT 1306
400 TyrAsnGluIleIleHisLeuMetGlnGln 409
1307 TATATATAAATTTATTAATCTAATGAGGAA 1336

```

seq_name: gb_pr:HUMILPCHL

seq_documentation_block:

```

LOCUS       HUMILPCHL             2493 bp             mRNA             PRI
DEFINITION  Human lysosomal acid lipase/cholesterol esterase mRNA, complete cds.
ACCESSION   M74775.1
VERSION     M74775.1 GI:187151
KEYWORDS    lysosomal acid lipase/cholesterol esterase.
SOURCE      Homo sapiens fibroblast cDNA to mRNA.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 2493)
AUTHORS     Anderson, R.A. and Sando, G.N.
TITLE       Cloning and expression of cDNA encoding human lysosomal acid
            lipase/cholesterol ester hydrolase. Similarities to gastric and
            lingual lipases
JOURNAL     J. Biol. Chem. 266 (33), 22479-22484 (1991)
MEDLINE     92042192
FEATURES
  source
    location/Qualifiers
      1..2493
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /tissue_type="fibroblast"
        /map="10"
        <1..2493
        /gene="LIPA"
        /note="G00-120-153"
        1..2493
        /gene="LIPA"
        41..1240
        /gene="LIPA"
        /BC_number="3.1.1.13"
        /note="alternate initiator ATG may be at position 47"
        /codon_start=1
        /db_xref="GDB:G00-120-153"
        /product="Lysosomal acid lipase/cholesterol esterase"
        /protein_id="AAA59519.1"
        /db_xref="GI:187152"

```


VERSION BC012287.1 GI:15126726
 KEYWORDS MGC.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2586)
 AUTHORS Strausberg, R.
 JOURNAL Direct Submission
 Submitted (06-AUG-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: villalona@bcm.tmc.edu
 Villalona, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
 A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
 Munzy, D.M., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 3 Row: 3 Column: 3
 This clone was selected for 'full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 434305.
 Location/Qualifiers
 1. 2586
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:5229 IMAGE:2900168"
 /tissue_type="Placenta, choriocarcinoma"
 /clone_id="NH_MGC_10"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"
 101. 1300
 /codon_start=1
 /product="Similar to lipase A, lysosomal acid, cholesterol
 esterase (Wolman disease)"
 /protein_id="AAH12287.1"
 /db_xref="GI:15126727"
 /translation="MKRRFLGLVVCVLWPLHSESGSGKLTALDPETNNMVSLEISYW
 GPSEELVETEDGYILCLNRIPGRKNSHSDGPKPVFLQHGILADSMMWYTLANS
 SLFLIADAGFVWGMGNSRGNTWSRKRTLSVSDDEFWAFSDMAKYLDPASINFTLL
 NKGQEOVVYVHGSGGTIGEIFASQIPELAKRTKMFALGPAVSAVAFSPAKKGR
 LPHRLIKLFGDEKRELPQSAFLKMGCHGCHTHYLKLCNLCITLGCPRERNLMSR
 VDYTTTSPAGTSYQNNLHNSQAVKRFKFAFDGSSAKNFTHTNOSTPTYNKDL
 VPIRVWSSGHDMLADYDVNILLQITNIVHESIPMEHLDFITWGLDAPWRYNKKIIL
 NMRKRG"
 BASE COUNT 736 a 511 c 531 g 808 t
 ORIGIN
 alignment_scores:
 Quality: 1303.00 Length: 410
 Ratio: 3.832 Gaps: 2
 Percent Similarity: 82.927 Percent Identity: 60.488
 alignment_block:
 US-09-578-063-47 x BC012287 ..
 Align seg 1/1 to: BC012287 from: 1 to: 2586
 2 LengUThTleuSeArGlnTrpIleValSerHsArGmeClumetr 18
 |||||||||||| || |||||.....|||
 CTCGAGACGACGCGCGCGGAG.....GACACCTCCAGATGATAAATGCG 111

[illegible]

193 heilealapheserthmetprogluleualaglnlylleysmettyr 209
 575 TTATAGATATTTTACAGATCCCTGAGCTGCTAAAGAGATTAATAATGTTT
 210 Phealaleualaproilealathrvallyshisalalysserprogllyth 226
 625 TTTCCTGGGCTGCTGGCTTCCGCTTCCGCTTCTGACTAGCCCTATGCG 674
 226 rlyspheleuleuleuproaspmetllysglyleiphegllyst 243
 675 CAATTTAGACAGATTACAGATCATCTCAATTAAGACTTATTGGAGACA 724
 243 ysglupheleutyrglnthrArpheleuaraglnleuVallyleu 259
 725 AAGATTCTTCCCGAGTGGCTTTTGAAGTGGCTGACCCAGCTT 774
 260 CysgllycInVallyleuAAspglnlleCysseranillemetleule 276
 775 TGCACTCATGTCTACTGAAAGAGCTGTGGAAATCTGTTTCTTCT 824
 276 ugllylPheasnThrAsnmetAsmetSerArGAlaserValtyra 293
 825 GTGGGATTTAATGAGAAATTAATATGCTAGAGTGAATATATA 874
 293 laalalsthrleuAlaglyThrservalGlnsnilleuHstPsr 309
 875 CAACACATCTCCCTGGAGAACTTCTGCAAAACATGTTACACTGGAGC 924
 310 GlAlalalAsnserglyleuleuArAlaphesprpGlysergluth 326
 925 CAGGCTGTAAATTCACAAAGTTTCAAGCTTGTACTGGGAGACCACTGC 974
 326 rlyAsnleuGlylucysAsnGlnProthrProvalArgtyrArgVala 343
 975 CAGAAATATTTCATTACACAGAGACTTATCCCTCCACATACATGTA 1024
 343 rGAspMetThrValProthrAlametryrphrglyglInasPTrpleu 359
 1025 AGGACATGCTTGGCCGACGCTGAGCGGGGCTCAGACTGCTT 1074
 360 SerAsnProGluAspVallysmetleuleuSerGluValthrAnleu 376
 1075 GCGAGATGCTACGACGTCATATCTTACTGATCAGATCAACCACTGCT 1124
 376 eTyHslAsnleuProglutThrAlaHsValasPhealjetrPGL 393
 1125 GTTCATGACAGACATTCGGAATGGAGCATCTGACTTCTGAGGCGC 1174
 393 euAspAlaProHsArgmettyrAsnGluIlelleHsleuMetGln 409
 1175 TGGATGCCCTTGGAGGCTTATTAATAATTAATTAATTAATGAGAAA 1224
 seq_name: gp_pr:HSU08464

seq_documentation_block: 2481 bp mRNA PRI 23-JUN-1994
 LOCUS HSU08464 2481 bp mRNA PRI 23-JUN-1994
 DEFINITION Human lysosomal acid lipase mRNA, complete cds.
 ACCESSION U08464
 VERSION U08464.1 GI:505052
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 2481)
 AUTHORS Du H. and Gregory G.A.
 TITLE Structural Conservation of Putative Functional Motifs between Mouse
 and Human Lysosomal Acid Lipase
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2481)
 AUTHORS Du H.
 TITLE Direct Submission

JOURNAL Submitted (08-APR-1994) Hong Du, Division of Human Genetics,
 Children's Hospital Medical Center, 3333 Bernet Street, Cincinnati,
 OH 45229-3039, USA
 FEATURES
 source
 1. 2481
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="clone hUAL"
 /clone_lib="HepG2 cDNA library/lambda Zap"
 /cell_line="HepG2"
 /cell_type="hepatoma"
 /tissue_type="liver"
 1. 33
 34. 96
 34. 1233
 /codon_start=1
 /product="lysosomal acid lipase"
 /protein_id="AAB60328.1"
 /db_xref="GI:505053"
 /translation="MKMRPLGLVCLVLTWLHSEGRKLTAVDEPTNNVSEIISY
 GPSEYIVETEDGYILCLNRIPIGRKNSDGPVPLQLGLADSSNWTNLANS
 SLGFIADGFPVMGNSRNTWSRKHTLSYODEFMAFSYDEMAKYDLPASINFI
 NKTGOEYVYVGHSGGTIGFLAFSOLPELARKIKMFALGPVAVAFCTSPMAKGR
 LPDLHLKDLFGRKELPOSAPLAKMLGTHCTHYILKELCNLCFLCGENENLNKR
 VDVTTHSPAGTSVQNMHMSQAVKRFQKQAFDWSSAKNTHINOSTPTTAVKQML
 VPTAVMSGHDLADVVDVNIILDTQNLVFHESIPWEHDLFIWGLDAPWRLYNKI
 NMRYQ"
 79
 variation
 /replace="c"
 97. 1230
 /product="lysosomal acid lipase"
 100
 variation
 /replace="g"
 1234. 2480
 2459. 2464
 2480
 3' UTR
 polyA_signal
 polyA_site
 BASE COUNT 692 a 477 c 513 g 799 t
 ORIGIN
 alignment_scores:
 Quality: 1298.00 Length: 400
 Ratio: 3.875 Gaps: 1
 Percent Similarity: 83.750 Percent Identity: 61.000
 alignment_block:
 US-09-578-063-47 x HSU08464 ..
 Align seq 1/1 to: HSU08464 from: 1 to: 2481
 12 SerHisArgMetGluMetTrpLeuIleValAlaTyrMetPheG1 28
 ||| |||||:|||||:||||| |||||:|||||:|||||
 25 AGCTCCAGATGAATAATCGCGTCTTGGGTTGGTGGTCTTGGTCT 74
 28 nArgAsnValAsnSerValHisMetProthrLys.....AlaValasp 43
 :|||:|||||:|||||:|||||:|||||:|||||:|||||
 75 CTGACCCCTCATCTAGAGGGGTCTAGAGGGAACCTGACAGCTGTGATC 124
 43 rGluAlaPheMetAsnIleSerGluIleIleGlnHisGlnGlyPro 59
 |||||:|||||:|||||:|||||:|||||:|||||:|||||
 125 CTGAACCAACATGATGTGAGTGAATAATTAATCTTACTGGGGAATCCCT 174
 60 CysGluGluTyrGluValAlaThrGluAspGlyTyrIleLeuSerValas 76
 |||||:|||||:|||||:|||||:|||||:|||||:|||||
 175 AGTAGAGAAATACCTAGTAGACAGAGAGATGATATCTGTGCTTAA 224
 76 nArgIleProArgGlyLeuValGlnProLysLysThrGlySerArgProv 93
 |||||:|||||:|||||:|||||:|||||:|||||:|||||
 225 CCGAATTCCTCATGGGAGAGAACCATTCGACAAAGGATCCCAACAG 274
 93 alValleuleuGlnHisGlyLeuValGlyValAlaSerAsnTrpIleSer 109
 |||||:|||||:|||||:|||||:|||||:|||||:|||||
 275 TTGCTTCTGCAACATGCTGCTGCGCAGATTCTAGTAAGTGGTCA 324

```

seq_name: gb_to:S81497
seq_documentation_block: 3144 bp mRNA ROD 02-AUG-2000
LOCUS S81497 lysosomal acid lipase-intracellular hydrolase [rats, Wolman, liver, RNA, 3144 nt].
ACCESSION S81497
VERSION S81497.1 GI:1336725
KEYWORDS Rattus sp., liver Wolman.
SOURCE Rattus sp.
ORGANISM Rattus sp.
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 3144)
Nakagawa,H., Matsudaira,S., Kuriyama,M., Yoshidome,H., Fujiyama,J., Yoshida,H. and Osame,M.
TITLE Cloning of rat lysosomal acid lipase cDNA and identification of the mutation in the rat model of Wolman's disease
J Lipid Res. 36 (10), 2212-2218 (1995)
JOURNAL 96129534
MEDLINE GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 176569] from the original journal article.
REMARK This sequence comes from Fig. 2.
FEATURES
SOURCE location/Qualifiers
1..3144 /organism="Rattus sp."
/db_xref="taxon:10118"
1..3144 /gene="lysosomal acid lipase, LAL"
233..1426 /gene="lysosomal acid lipase, LAL"
/gene="lysosomal acid lipase, LAL"
/note="intracellular hydrolase; This sequence comes from Fig. 2; conceptual translation presented here differs from translation in publication; LAL"
/codon_start=1
/product="lysosomal acid lipase"
/protein_id="AAB36043.2"
/db_xref="GI:3653291"
/translat="MOLLGRVYICFVVGILLSGPRTGISAVDPANNWVEIIMHGGTPEHVSTGTDGVILGVHRIPHGKRNQPKGPVVYLQMRHQFLADSSNWVINIDNSIGFIILADAGFDVMGNSRGNTWSRKHKTLISVSODEVAFSFDEMAKYDLPASINYLNRKTGOGLLVNVHSOGCTGFIASFOMPELAKVKKEFPALAPVLSNFASGPVAKGRLEDLLEDLFEGOKPLPOSAMVWLSTFHICTHYIMEICANIFELCGFNENKLNNSRDVTYTHHCAGTSVQNMMHWTVYVKKHLQAEDPMGSDDKNFRITNGSYPLPSIKRMQPLTALWSGKWLDLRSDINILTEPTLTVPYHNKIPEWHLDFTWGLDAPWRLYNEVSLMKTKO"
BASE COUNT 852 a 690 c 754 g 848 t
ORIGIN
alignment_scores:
Quality: 1220.50 Length: 378
Ratio: 3.838 Gaps: 3
Percent Similarity: 84.127 Percent Identity: 58.201
alignment_block:
US-09-578-063-47 x S81497 ..
Align seg 1/1 to: S81497 from: 1 to: 3144
37 Prothr.....Lysalavalaspproglualalaphemetarnlesle 50
||||| :::||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
290 CCACGCCGGACCATTATCAGCTGTGGACCCCTGAAGCAAAATGAAATGTGAC 339
50 rgluilelleglnhslnglytyrProcysegiugljutyrgluvalalat 67
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
340 TGAAGATAATCATGCACACTGGGATATCCA.....GAGCACTAGTACAGA 383
67 hgcluaspoljyttrileusevalasnargljetproargelyleuVal 83
||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
384 -CAGGAATGGCTACATTCGGTGTTCCACCGAATCCCGATGGCGGGAAG 433

```



```

292 yrlaiaiahstlhleuadlaglythrservalglnasnlleuwnsttr 308
      :::::::::::::: ||| ||::::::::::::::::::|||
851 atgaaacacacgtccctgcggaaacttctgcacaaacatgtttacactgg 900
      :::::::::::::: ::::::::::::::::::::|||
309 serglnalavalasnserygluleuavalapheasprtrpglysergl 325
      :::::::::::::: ::::::::::::::::::::|||
325 utlthlyasnlleuglulyscysasnglnprthrprovalargylargv 342
      :::::::::::::: ||||| ||||| ||| ::::::::::
951 tgaamaaacaattatttttcattacacacagagttttctccctgcgacacaa 1000
342 alagaspmetlnrhalprothralametrpthngllyglinalasprtr 358
      :::::::::::::: ::::::::::::::::::::|||
1001 taataaaacatgcgtcttaccactgctttttggagcggggccggagactgg 1050
359 leuSERasnproglnuaspvalylusmetleuLeuSERglnuvalthrAsnLe 375
      :::::::::::::: ::::::::::::::::::::|||
1051 ctgcacagacatcaatgacatcatcttaccagcttcaacgtccaatcccaagtt 1100
375 uilctytlhlslyasnlleproglulnrplahlsvalasphellettprg 392
      :::::::::::::: ::::::::::::::::::::|||
1101 gggtattcacaaagacattctcgtagggagacatctggaattttttatctggc 1150
392 lyleuAspAlaProHlsArgMetTyrAsnGluIleIleHlsleuMetGln 408
      :::::::::::::: ::::::::::::::::::::|||
1151 ctttgagatgcccccctggaacactgtatgaaatgaaatattacgtctatagag 1200
409 glnGlnGlu.....ThrAsnLeuSERglnGly 417
      :::::::::::::: ||::::::::::::::::::||
1201 aaataccagtgaaatgctgcagcttgaaatgacttcccttcacagagggg 1245

seq_name: gb_pat:A26689

seq_documentation_block:
LOCUS      A26689                1378 bp    DNA                    PAT
DEFINITION Precursor of rabbit gastric lipase coding sequence.
ACCESSION  A26689
VERSION    A26689.1  GI:905029
KEYWORDS
SOURCE     synthetic construct.
ORGANISM   synthetic construct
            artificial sequence.
REFERENCE  1 (bases 1 to 1378)
AUTHORS    Benicourt,C., Blanchard,C. and Junien,J.L.
TITLE      Recombinant gastric lipase from rabbit and pharmaceutical
            compositions
            Patent: EP 0542629-A 9 19-MAY-1993;
            INSTITUT DE RECHERCHE JOUVEINAL
FEATURES
            Location/Qualifiers
            source          1..1378
                           /organism="synthetic construct"
                           /db_xref="taxon:32630"
BASE COUNT      397 a      287 c      276 g      418 t
ORIGIN
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832

```

```

98 AACACATGCTGCTTTTGGAAATGACACACA.....AACCC 138
43 rogluAlaPheMetAsnIleSerGluIleIleGlnHisGlnGlyPro 59
139 CTGAAGTAATATGAAATATATAGTACAGATGATTCCTACTGGGACCA 188
60 CysGluGluTyrGluValAlaThrGlnAspGlyTyrIleLeuSerValAs 76
189 AGTGAATAATATGAAGTTGTCAGTGAAGATGTTATCTTGAAGCTTAA 238
76 nargIleProArgGlyLeuValGlnProLysThrGlySerArgProV 93
239 CAGAAATTCCTATGAGGAAGAAATTCAGGAACAGAGCCAGAGACCCG 288
93 alValLeuLeuGlnHisGlyLeuValGlyGlyAlaSerAsnTrpIleSer 109
289 TCGTAATTTTGCAGCATGGTTGCTTCATCAGCCTCAACTGATCTCC 338
110 AsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAspAlaGlyPheAs 126
339 AACCTGCCCAACACAGCTGCGCTTATTTCTGGCAGATGCTGTTATGG 388
126 pValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerArgLysHisLysT 143
389 TGTGTGCGCTGGGAACAGCAGAGAAATACCTGCTCCAGAAAGAACTTAT 438
143 hrLeuSerIleAspGlnAspGluPheTrpAlaPheSerTyrAspGluMet 159
439 ACTATTCACAGACCTCCCTGAATTCCTGGCTTTTACCTTGATGATAATG 488
160 AlaArgPheAspLeuProAlaValIleAsnPheIleLeuGlnLysThrGln 176
489 GCTAAATGTGACCTTCAGCCACCAATTCATTCATTTGAAGAAACGTGG 538
176 yGlnGluLysIleTyrValGlyTyrSerGlnGlyThrTrpMetGlyP 193
539 ACAGAAAGAGCTGCATATGTTGTCATCCAGGACACACCATGTT 588
193 heIleAlaPheSerThrMetProGluLeuAlaGlnLysIleLysMetYr 209
589 TTATTCGCTTTTCTCAAAATCCCAAGCTGGCTGAAGAATCAAAACCTTT 638
210 PheAlaLeuAlaProIleAlaThrValLysHisAlaLysSerProGlyTh 226
639 TATGATTAAGTCCAGTCGACCTGGAATACAAAAAGCCTGTAA 688
226 rLysPheLeuLeuLeuProAspMetMetIleLysGlyLeuPheGlyLysL 243
689 CAACCTTAAGTTATTCCTCCACACCATGTCAGATTTATTTGGTGACA 738
243 ysgLupheLeuTyrGlnThrArgPheLeuArgGlnLeuValIleTyrLeu 259
739 AAATATTCACCCACACAAATTCCTTTCATCAATTCCTGCACATCAAGTG 788
260 CysGlyGlnValIleLeuAspGlnIleCysSerAsnIleMetLeuLeu 276
789 TGTTCGCCGTGAGACACTGAATGATTAATTTGCAGCAATGCCCTTAATAT 838
276 uGlyGlyPheAsnThrAsnAsnMetLysMetSerArgAlaSerValTyrA 293
839 TTGTGGCTTTGACAGCGCAACTGGAACATGATGCTGGATGTGTACG 888
293 laAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeuHisTrpSer 309
889 TGTCAATCAATATCCGACAGAACTTCAGTTCAAAACATGCTGCACATGAC 938
310 GlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpGlySerGluTh 326
939 CAGGCTGTTAAATCTGGAATTTTCACGCTTTTAATTTGGGGAAGCCAGC 988
326 rLysAsnLeuGluLysCysAsnGlnProThrProValArgTyrArgValA 343

```

```

989 TCAGATGTAGTGCACCTCAATTCAGCCACACCTCCCTACTACATGTGA 1038
343 rgaAspMetThrValProThrAlaMetTrpThrGlnGlyGlnAspTrpLeu 359
1039 CCGCCATGAATGTGCATTAATGAGTGTGAGTGTGGCAATGACTGGTTG 1088
360 SerAsnProGluAspValLysMetLeuLeuSerGluValAlaThrAsnLeu 376
1089 GGTGACCCCAAGATGTGACCTTTTGGCTTCCAAAACTTCAATCTCAT 1138
376 eTyrHisLysAsnIleProGluTrpAlaHisValAspPheIleTrpGlyL 393
1139 TTACCAACAGAGATTTCTTCATATTAATCACTTGATTTATCTGGCCAA 1188
393 euAspAlaProHisArgMetTyrAsnGlnIleIleHisLeuMetGlnGln 409
1189 TGAATGCTCTTCAGAAAGTTTACAAATGAATTAATTTCTATGATGCCAAA 1238
410 GluGlu 411
1239 GATAAA 1244

```

```

seq_name: gb_pat:A26690
seq_documentation_block: 1378 bp DNA PAT 05-APR-1995
LOCUS A26690
DEFINITION Precursor of rabbit gastric lipase (comp.).
ACCESSION A26690
VERSION A26690.1 GI:905030
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
INSTITUT DE RECHERCHE JOUVEINAL

```

```

FEATURES
source
BASE COUNT 418 a 276 c 287 g 397 t
ORIGIN

```

```

alignment_scores:
Quality: 1167.00 Length: 402
Ratio: 3.505 Gaps: 3
Percent Similarity: 82.856 Percent Identity: 54.478

```

```

alignment_block:
US-09-578-063-47 x A26690/rev ..

```

```

Align seg 1/1 to reverse of: A26690 from: 1 to: 1378

```

```

16 GluMetTrpLeuLeuIleLeuValAlaTyrMetPheGlnArgAsnValAs 32
1328 AAAATGTGGTGCCTTTTCATGATGTTGACACTTGTGATCTGCA...CTTGG 1282
32 nSerValHis.....MetProThrLysAlaValAsp 43
1281 AACATACATGCTGCTTTTGGAAATGACACACA.....AACCC 1241
43 rogluAlaPheMetAsnIleSerGluIleIleGlnHisGlnGlyPro 59
1240 CTGAAGTAATATGATATATAGTACAGATGATTCCTACTGGGGAATACCA 1191
60 CysGluGluTyrGluValAlaThrGlnAspGlyTyrIleLeuSerValAs 76
1190 AGTGAATAATATGAAGTTGTCAGTGAAGATGTTATATCTTGAAGCTTAA 1141

```

```

76  nAgtIleProArgGlyLeuValGlnProLysLysThrGlySerArgProV 93
   |||||
1140 CAGAAATTCCTATGCGAAGAAATTCAGGAAACAGAGCCAGAGACCCG 1091
   |||||
93  aValLeuLeuGlnHisGlyLeuValGlyGlyAlaSerAsnTrpIleSer 109
   |||||
1090 TCGATATTTTGGCAGCAGATGTTGGCTTCATCAGCTCAAACTGATCTCC 1041
   |||||
110  AsnLeuProAsnSerLeuGlyPheIleLeuAlaSerAlaGlyPheAs 126
   |||||
1040 AACCTGCCCAACACAGCCTGGCTTATCTGCGAAGTGGTGTATGG 991
   |||||
126  pValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerArgLysHisLys 143
   |||||
990 TGTGTGCTGGGAGACAGCAGAGAAATACCTGCTCCAGAAAGAACTTAT 941
   |||||
143  hIleuSerIleAspGlnAspGlnPheTrpAlaPheSerTrpAspGluMet 159
   |||||
940 ACTATTCACAGACTCCGTTGAATTCGGCTTTACCTTGATGAATG 891
   |||||
160  AlaArgPheAspLeuProAlaValIleAsnPheIleLeuGlnLysThr 176
   |||||
890 GCTAAATATGACCTCCAGCACAATGACTTCATGTAAAGCAACTGG 841
   |||||
176  yGlnGluLysIleTyrTyrValGlyTyrSerGlnGlyThrThrMetGly 193
   |||||
840 ACAAGAAACCTGCACTATGTGTGTCATCTCCAGGCAACCCATTGTT 791
   |||||
193  hIleAlaPheSerThrMetProGluLeuAlaGlnLysIleLysMetGly 209
   |||||
790 TTTATGCTTTTCTACAAATCCCAAGCTGGCTGAAGATCAAAACCTTT 741
   |||||
210  PheAlaLeuAlaProIleAlaThrValLysHisAlaLysSerProGly 226
   |||||
740 TATGCAATACCTCCAGTCCGACACTGTAAGTATCAAAACCTGTGA 691
   |||||
226  rLysPheLeuLeuProAspMetMetIleLysGlyLeuPheGlyLysL 243
   |||||
690 CAAATTCAGTTTATCTCCCAACCATGTCACAAATATATATGATGTC 641
   |||||
243  yGlnLeuLeuTyrGlnThrArgPheLeuArgGlnLeuValIleTyrLeu 259
   |||||
640 AAATATCTCAACCCACAAATTTCTGATCAATCTCTCCACACACAGTG 591
   |||||
260  CysGlyLeuValIleLeuAspGlnIleCysSerAsnIleMetLysLeu 276
   |||||
590 TGTCCCGTGAACACACTGAATGTAATTCGCAATGCTCATTTATTA 541
   |||||
276  uGlyGlyPheAsnThrAsnMetAsnMetSerArgAlaSerValTyr 293
   |||||
540 TTGTGCTTTGACAGCGCAATCTGAACATGAGTGGCTGGATGTATG 491
   |||||
293  IaaIleThrLeuAlaGlyThrSerValGlnAsnIleLeuHisTrpSer 309
   |||||
490 TGTCACTAATCCGCGAGCAATCTGATCAAAACATGCTCAGTGGAG 441
   |||||
310  GlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpGlySerGly 326
   |||||
440 CAGGCTGTGAATCTGGAATTTTCAAGCTTTAATTCGGGAACCTGCA 391
   |||||
326  rLysAsnLeuGluLysCysAsnGlnProThrProValArgTyrArgVal 343
   |||||
390 TCAGAAATGATGTCATTCATACGCCACACCTCCCTACTACAAATGTA 341
   |||||
343  rGAspMetThrValProThrAlaMetTrpThrGlyGlnAspTrpLeu 359
   |||||
340 CCGCATGATGTCGCAATGTCAGTGTGAGTGGCAATGACTGGTGG 291
   |||||
360  SerAsnProGluAspValLysMetLeuLeuSerGluValThrAsnLeu 376
   |||||
290 GCTGACCCCAAGATGTTGACTTGTCTGCAAACTTTCTAATCTCAT 241
   |||||
376  eTyrHisLysAsnIleProGluTrpAlaHisValAspPheIleTrpGly 393

```

```

|||||
240  TTACCACAGAGAGATCTTCATATATACACTGGAATTTATCTGGGCAA 191
   |||||
393  euAspAlaProHisArgMetTyrAsnGlnIleIleHisLeuMetGln 409
   |||||
190  TGAATGCTCCCAAGATTTAACAATGAATTAATTTCTATGATGCAAAA 141
   |||||
410  GluGlu 411
   |||||
140  GATAAA 135
   |||||

seq_name: gb_pat:A01157
seq_documentation block: 1336 bp mRNA PAT 29-JAN-1993
LOCUS A01157
DEFINITION R.norvegicus mRNA for prelingual lipase protein.
ACCESSION A01157
VERSION A01157.1 GI:14748
KEYWORDS lipase.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1336)
AUTHORS Williamson,R.
JOURNAL Patent: WO 8500381-A 3 31-JAN-1985;
Celltech Ltd
FEATURES
source location/Qualifiers
1..1336
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
7..1194
/codon_start=1
/product="prelingual lipase"
/protein_id="CAA00136.1"
/db_xref="GI:14749"
/db_xref="SWISS-PROT:P04634"
/translation="MVLITIVISTFGGAGHLFGKLGKRGNEPANKNISQMTTYGCP
COEYEVTEIDGYTLGVRIPIHKNNSENISGKRPVYIQLGLIASATNIAINLPNNSLA
FMIAADGYDWLWNSKRNWISPDSEVFAFSDENAKYDLPTNIFIVQKT
GOEKIYVGHSGQTTIGFAFSNPTLAKIKFYALAVATKYTQSLKISFPT
FLPKLMFGKMKPLPHYTPDDPLGTGECSEVLDLCSNLPFTFCGDKKLANVREDY
YLGHNPAQSYVDPLHMAOLVDSGKTFQAPNMSPSQNMILHYOKPPEVDGAMTPRV
AVWNGNDLADPDVAMLPKLSNLPFKELIATYINHLDFIAMDAPQEVYENHISM
ABD"
BASE COUNT 409 a 283 c 265 g 379 t
ORIGIN

alignment_scores:
Quality: 1166.50 Length: 399
Ratio: 3.556 Gaps: 1
Percent Similarity: 82.206 Percent Identity: 54.637

alignment block:
US-09-578-063-47 x A01157 ..
Align seg 1/1 to: A01157 from: 1 to: 1336

16  GluMetTrpLeuLeuIleLeuValAlaTyrMetPheGln...ArgAsnVa 31
   |||||
4  AAGATGTGGCTGTATTAATACAAAGTGTGATATCAACATTCGGAGGTGC 53
   |||||
31  IAsnSerValHisMetProThrLysAlaValAspProGluAlaPheMet 48
   |||||
54  ACATGGCTATATTGGAACACTGGGTCTCGGAACCCGGAACCAATATGTA 103
   |||||
48  snIleSerGluIleIleGlnHisGlnGlyTyrProCysGluGluTyrGlu 64
   |||||
104  ATATTAGCAGATGATTAACCTTACGTGGGATATTCATGTCAAGAAATATGA 153
   |||||
65  ValAlaThrGluAspGlyTyrIleLeuSerValAsnArgIleProArgGly 81

```

```

1054 GTCGGCAGCTGCTCCCAAACTCCCAACTCCTGTTCCATAAGAGAT 1103
381 eProglutrinAlaHisValAspPheIleTyrPcyluAspAlaProHisA 398
1104 TCTTCCTCCATCATCACTGAGCTTCACTGGGCAATGATGCTCCCAAG 1153
398 rgMetTyrAsnGluIleIleHisLeuMetGlnGluGluIurHsrn 413
1154 AGGTTTACATGATGATGATTTCATGATGCGAGAACTAATAAGANT 1200
seq_name: 'gb.ro:RNLIP'

seq.documentation.block:
LOCUS      RNLIP               1355 bp      mRNA           ROD           30-MAR-1995
DEFINITION Rat mRNA for lingual lipase..
ACCESSION  X02309
VERSION    X02309.1 GI:56595
KEYWORDS   glycoprotein; lipase.
SOURCE     Norway rat.
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE  1 (bases 1 to 1355)
AUTHORS   Docherty,A.J., Bodmer,M.W., Angal,S., Verger,R., Riviere,C.,
            Lowe,P.A., Lyons,A., Emlage,J.S. and Harris,T.J.
TITLE     Molecular cloning and nucleotide sequence of rat lingual lipase
            cDNA
JOURNAL    Nucleic Acids Res. 13 (6), 1891-1903 (1985)
MEDLINE    85215587
COMMENT    Data kindly reviewed (12-FEB-1986) by A.J.P. Docherty.
FEATURES
            source
                1..1355
                /organism="Rattus norvegicus"
                /db_xref="taxon:10116"
                sig_peptide
                    49..102
                    /product="put. signal peptide (aa -18 to -1)"
                    49..1236
                    /note="lipase precursor"
                    /codon_start=1
                    /protein_id="CAA26179.1"
                    /db_xref="GI:56596"
                    mat_peptide
                        103..1233
                        /product="lipase (aa 1-377)"
                        505..507
                        /note="aa 153 pot. lipid binding site"
                        misc_feature
                            1355
                            /note="polya site"
BASE COUNT  400 a 291 c 276 g 387 t 1 others
ORIGIN
alignment_scores:
    Quality: 1166.50      Length: 399
    Ratio: 3.556          Gaps: 1
Percent Similarity: 82.206 Percent Identity: 54.637

alignment_block:
US-09-578-063-47 x RNLIP ..
Align seg 1/1 to: RNLIP from: 1 to: 1355
16 GUGUETTPLEUULEUULEUVALAATYRMECPHEGHN..ARGASNYA 31
:::

```

seq_name: gb_om:CFLIPASE

seq_documentation_block: 1651 bp mRNA MAM 18-JUN-1998

LOCUS CFLIPASE Canis familiaris mRNA for lipase.

DEFINITION Y13899 GI:2204112

VERSION Y13899.1

KEYWORDS DGL gene; gastric lipase; triacylglycerol lipase.

SOURCE dog.

ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 1651)
Vaganay,S., Jolliff,G., Bertaux,O., Toselli,E., Devignes,M.D. and Benicourt,C.
The complete cDNA sequence encoding dog gastric lipase
DNA Seq. 6 (4), 257-262 (1998)

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source location/Qualifiers

1..1651 /organism="Canis familiaris"
/db_xref="taxon:9615"
/dev_stage="adult"
/tissue_type="stomach mucosa"
54..1250 /gene="DGL"
54..1250 /gene="DGL"
/EC_number="3.1.1.3"
/codon_start=1
/product="triacylglycerol lipase"
/protein_id="CAA74198.1"
/db_xref="GI:2204113"
/db_xref="SWISS-PROT:P80035"
/translation="MWLLPLPAASYISTLTGTHGLFGRKLHPTNPEVTNNISOMITTYMGY
PAAEEYEVTEDEGYILGIDRIPIYGRKNSENIGRVAALQHGILASATNMWISNPNSNL
AFILADGDYDWLGNBSRGMPARNRLKYSPDSVEFMASFPEMAKYDLPATIDEILKKL
TGDKRLTVGHSSOSTYTGFATFSNPRLARKIFVALAPATVKYTEILLIKMLIVP
SFELKLFGRNKLETFPHHFDFDLATEVRESREVDLLCSNALFTICGFDTMKLNRSD
VYLSEHPAGTSVQNVQLHWQAQVKSQGKAQDAQDGSPVQNMHHYHQSNMPYYNLDMHVAP
IAVNGENDLLADHDVDDLILSKLPNIHYRKIPRYNHLDIFIMADAPQAVYNIEIVSM
MGTDNKK"

sig_peptide 54..110
/gene="DGL"
/evidence-experimental

mat_peptide 111..1247
/gene="DGL"
polya_signal 1602..1607
BASE COUNT 432 a 419 c 352 g 448 t
ORIGIN

alignment_scores:
Quality: 1154.00 Length: 398
Ratio: 3.486 Gaps: 5
Percent similarity: 83.166 Percent identity: 54.271

alignment_block:
US-09-578-063-47 x CFLIPASE ..
Align seg 1/1 to: CFLIPASE from: 1 to: 1651

```

17 MetTrpLeuLeuIleValAlaIleValMetPheGlnArgAsnValAsn 33
|||||
54 ATGTGGCTGCTTTAAACAGCGCGAGTGTGATCT...ACACTGGAAAC 100
33 rValHis.....MetProThrLysAlaValAspProG 44
|||||
101 CACACATGGTTATTGGAAATATACATCCACA.....AACCTG 141
44 IuAlaPheMetAsnIleSerGluIleIleGlnHisGlnGlyTyrProCys 60
|||||
142 AAGTACACATGAAATATAGTACAGATGATCCTACTGGGATACCACT 191
61 GluGluTyrGluValAlaIleThrGlnAspGlyTyrIleLeuSerValAsn 77
|||||
192 GAGGAAATAGAGTGTGACCGAAGACGGTATATCTTGGGATCGACAG 241
77 GlieProArgGlyLeuValGlnProLysLysThrGlySerArgProVal 94
|||||
242 AATTCCTATGGAGAGAAATATTCAGAGATATAGCCGAGACCTGTGG 291
94 aIleuLeuGlnHisGlyLeuValGlyGlyAlaSerAsnTrpIleSerAsn 110
|||||
292 CATTTTGGACACCGTTTGGCTGCATCAGCCACAACATGATCTCCAAC 341
111 LeuProAsnAsnSerLeuGlyPheIleLeuAlaAspAlaGlyPheAsp 127
|||||
342 CTGCCCAACACAGCTGCGCTTCATCCGCGGAGCGCGGTAGACGT 391
127 ITrpMetGlyAsnSerArgGlyAsnAlaTrpSerArgLysHisLysThr 144
|||||
392 GTGGCTGGGAGACAGAGCGGACACACCTGGCGGAGAGATCTGTACT 441
144 euSerIleAspGlnAspGluPheTrpAlaPheSerTyrAspGluMetAla 160
|||||
442 ACTCCCGGACCTCGCATATCTGGCTTTCAGCTTGACAGATGGCT 491
161 ArgPheAspLeuProAlaValIleAsnPheIleLeuGlnLysThrGly 177
|||||
492 AATATGACCTTCGCCACCATGTACTCTTGAAGAAACGGGAGAA 541
177 nGluLysIleTyrValGlyTyrSerGlnGlyThrMetGlyPheI 194
|||||
542 GGACACAGACACTAGTGGCTTCCAGGACACACCATGTTGTTCA 591
194 IeAlaPheSerThrMetProGluLeuAlaGlnLysIleLysMetTyrPhe 210
|||||
592 TCGCCTTTTCCACCAATCCAGCTGCGGAAACGATCAAAACCTTCTAT 641
211 AlaIleuAlaProIleAlaThrValLysHisAlaLysSerProGlyThr 227
|||||
642 GCATTAGCTCCCGTGGCCACCGTGAATACACGAAACCTTTAAACA 691
227 sPheLeuLeuLeuProAspMetMetIleLysGlyLeuPheGlyLysG 244
|||||
692 ACTCATGCTGCTGCTGCTCTCAAGCTTATATTGGAACAAANA 741

```

```

244 IupheLeuTyrGlnThrArgPheLeuArgGln...LeuValIleTyrLeu 259
|||||
742 TATTC...TACCCACACCACTTCTTGTGATCAATTTCTGCCACCAAGTA 788
260 CysGlyGlnValIleLeuAspGlnIleCysSerAsnIleMetLeuLeu 276
|||||
789 TGCTCCCGGAGAGCGGTGATCTCTTCGACAGACAGCGCTGTATCAT 838
276 uGlyGlyPheAsnThrAsnAsnMetAsnMetSerArgAlaSerValTyr 293
|||||
839 TTGTGATTTGACACTAGAACTGATGAACATGATGCTGGTGTGTATC 888
293 IeAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeuHisTrpSer 309
|||||
889 TGTCAATATTCACAGAGACATCGTTCCAGAACGTCTCCACTGTCTC 938
310 GlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpGlySerGlu 326
|||||
939 CAGCGCTTAAAGTCTGGGAAAGTCCAGACTTTGACTGGGAGACCACT 988
326 rLysAsnLeuGluLysCysAsnGlnProThrProValArgTyrArgVal 343
|||||
989 TCAGAAACATGATGACATCATCAGACATGCCCTCACTACAACTGA 1038
343 rGAspMetThrValProThrAlaMetTrpThrGlyGlyGlnAspTrpLeu 359
|||||
1039 CAGACATCATGTCGCAATCGCATGTGAGAGGAGGAGGAGCACTGGTG 1088
360 SerAsnProGluAspValLysMetLeuLeuSerGluValThrAsnLeu 376
|||||
1089 GCCGACCTCCAGATGTGACCTTTGCTTTCAGAGCTCCCAAACTCAT 1138
376 eTyrHisLysAsnIleProGluTrpAlaHisValAspPheIleTrpGly 393
|||||
1139 TTACCAAGAGAAATTCCTCCATCAATCAGCTTGATCTGATCTGGGCA 1188
393 euAspAlaProHisArgMetTyrAsnGluIleIleHisLeuMet 407
|||||
1189 TGGATGCCCTCCAGCGGTTTCAATGAATGATTTTCATGATG 1232

```

seq_name: gb_pat:AR039022

seq_documentation_block: 1137 bp DNA PAT 29-SEP-1999
LOCUS AR039022
DEFINITION Sequence 2 from patent US 5807726.
ACCESSION AR039022
VERSION AR039022.1 GI:5958385
KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1137)
AUTHORS Blanchard,C., Benicourt,C. and Junien,J.
TITLE Nucleic acids encoding dog gastric lipase and their use for the
JOURNAL production of polypeptides
PUBLISHED Patent: US 5807726-A 2,15-SEP-1998;
FEATUES Location/Qualifiers
source 1..1137

BASE COUNT 296 a 308 c 249 g 284 t
ORIGIN

alignment_scores:
Quality: 1141.50 Length: 372
Ratio: 3.624 Gaps: 3
Percent similarity: 84.677 Percent identity: 56.452

alignment_block:
US-09-578-063-47 x AR039022 ..
Align seg 1/1 to: AR039022 from: 1 to: 1137


```

37  ProthrlysaValaAspProgluAlaPheMetasniLeuSerGluIleI 53
    ||||| ..... ||||| ||||| |||||
19  CCCCA.....AACCTGAAGTGAACCTGATATATAGTCAAGAT 59
53  eGlnHisGlnGlyTrpProCysGluGluTrpGluValAlaIleThrGluAsp 70
    ||||| ..... ||||| ||||| ||||| |||||
60  CACTACTGGGGATACCCAGCTGAGGAATATGAGATTGTGACCAAGACG 109
70  lYTrIleLeuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
    ||||| ..... ||||| ||||| ||||| |||||
110 GTTATATCTTGGGATGACAGAAATTCCTTATGGAGGAAAAAATTCAGAG 159
87  LysThrGlySerArgProValIleLeuGlnIleHisGlyLeuValGlyI 103
    ||||| ..... ||||| ||||| ||||| |||||
160 AATATAGCGCGAGACCTGTTGCAATTTTGCACACGGTTGCTGCGATC 209
103 yAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleI 120
    ||||| ..... ||||| ||||| ||||| |||||
210 AGCCACAACCTGATCTCAACCTGCCCAACACAGCTGCTTCATATCC 259
120 euAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136
    ||||| ..... ||||| ||||| ||||| |||||
260 TGCGCGACGCGGCTAGACGCTGCGCTGGGAGACAGCAGGCGCAACACC 309
137 TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
    ||||| ..... ||||| ||||| ||||| |||||
310 TGGGCGACGAGGAATCTGACTACTCGCCGACCTCGCTGCTGCGC 359
153 aPheSerTrpArgLysMetAlaArgPheAspLeuProAlaValIleAsn 170
    ||||| ..... ||||| ||||| ||||| |||||
360 TTATAGCTTGTACGAGATGCTAATATGACTTCGCCCAACCATGACT 409
170 heIleLeuGlnLysThrGlyGlnGluLysIleTrpValGlyTrpSer 186
    ||||| ..... ||||| ||||| ||||| |||||
410 TCATCTTGAAGAAACGGAGACAGCAAGCTACACTAGCTGGCATTC 459
187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
    ||||| ..... ||||| ||||| ||||| |||||
460 CAGGCGACACCATGTTGCTTCATCGCTTTCACCAATCCCAAGCTGCG 509
203 agnLysIleLysMetTrpPheAlaLeuAlaProIleAlaThrValLysH 220
    ||||| ..... ||||| ||||| ||||| |||||
510 GAACGGATCAAAACCTTATGCAATGATCCGCTGCCACCGGAAGT 559
220 lsaLysSerProGlyThrLysPheLeuLeuLeuProAspMetMetIle 236
    ||||| ..... ||||| ||||| ||||| |||||
560 ACACCGAAACCTGTAAACAACATCATGCTGCTGCTTCTCTCTTC 609
237 LysGlyLeuPheGlyLysLysGluPheLeuTrpGlnThrArgPheLeuAr 253
    ||||| ..... ||||| ||||| ||||| |||||
610 AAGCTTATATTGGAAACAAATATTC...TACCACACACACTCTTGA 656
253 gGln...LeuValIleTrpLeuGlyGlnValIleLeuAspGlnIleC 269
    ||||| ..... ||||| ||||| ||||| |||||
657 TCAATTTCTCCGACCGAGTATGCTCCCGGAGACGGTGAATCTCTCT 706
269 ySerAsnIleMetLeuLeuGlnGlyPheAsnThrAsnAsnMetLys 285
    ||||| ..... ||||| ||||| ||||| |||||
707 GCACCAACGCCCTGTTATCATTTGTGATTTGACACATATGACTTGAAC 756
286 MetSerArgAlaSerValTrpAlaAlaHisThrLeuAlaGlyThrSerVa 302
    ||||| ..... ||||| ||||| ||||| |||||
757 ATGAGTCGCTTGGATGTGTATCTCATATATCCAGCAGACATCTCGT 806
302 lGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyLeuLeuArg 319
    ||||| ..... ||||| ||||| ||||| |||||
807 TCAGAACCTGCTCCACTGCTCCAGGCTGTAAGCTGGGAGATTCACAG 856
319 lApeAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 335
    ||||| ..... ||||| ||||| ||||| |||||
857 CTTTGTGCTGGGAGGCCAGTTCAGAACATGATGACTATCATCATCAGAC 906
336 ThrProValArgTrpArgValArgAspMetThrValProThrAlaMetTr 352

```

```

907 ATGCTCTCCCTACTACAACTGACAGACATGCATGCGCAATCGAGTGT 956
352 pThrGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeu 369
    ||||| ..... ||||| ||||| ||||| |||||
957 GACGCTGGCAACGACTGCTGCGCGACCCYACAGATGTGACCTTTGCG 1006
369 euSerGluValThrAsnLeuIleTrpHisLysAsnIleProGluTrpAla 385
    ||||| ..... ||||| ||||| ||||| |||||
1007 TTCCAGACTCCCAATCTCATTTACCAACAGGAAGATTCCTTCAAT 1056
386 HisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTrpAsn 402
    ||||| ..... ||||| ||||| ||||| |||||
1057 CACTTGGACTTATCTGGGCAATGATGCCCTCAACGGTTTCAATGA 1106
402 uIleIleHisLeuMet 407
    ||||| ..... ||||| ||||| ||||| |||||
1107 AATGTTCATGATG 1122

seq_name: gb-pat:AR092633

seq_documentation_block:
LOCUS AR092633 1137 bp DNA PAT 08-SEP-2000
DEFINITION Sequence 2 from patent US 5998189.
ACCESSION AR092633
VERSION AR092633.1 GI:10019385
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1137)
AUTHORS Blanchard,C., Benicourt,C. and Junten,J.
TITL E Polypeptide derivatives of dog gastric lipase and pharmaceutical
compositions containing same
JOURNAL Patent: US 5998189-A 2 07-DEC-1999;
FEATURES
location/Qualifiers
source 1..1137
BASE COUNT 296 a 308 c 249 g 284 t
ORIGIN

alignment_scores:
Quality: 1141.50 Length: 372
Ratio: 3.624 Gaps: 3
Percent Similarity: 84.677 Percent Identity: 56.452

alignment_block:
US-09-578-063-47 x AR092633 ..

Align seg 1/1 to: AR092633 from: 1 to: 1137

37 ProthrlysaValaAspProgluAlaPheMetasniLeuSerGluIleI 53
    ||||| ..... ||||| ||||| ||||| |||||
19 CCCCA.....AACCTGAAGTGAACCTGATATATAGTCAAGAT 59
53 eGlnHisGlnGlyTrpProCysGluGluTrpGluValAlaIleThrGluAsp 70
    ||||| ..... ||||| ||||| ||||| |||||
60 CACTACTGGGGATACCCAGCTGAGGAATATGAGATTGTGACCAAGACG 109
70 lYTrIleLeuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
    ||||| ..... ||||| ||||| ||||| |||||
110 GTTATATCTTGGGATGACAGAAATTCCTTATGGAGGAAAAAATTCAGAG 159
87 LysThrGlySerArgProValIleLeuGlnIleHisGlyLeuValGlyI 103
    ||||| ..... ||||| ||||| ||||| |||||
160 AATATAGCGCGAGACCTGTTGCAATTTTGCACACGGTTGCTGCGATC 209
103 yAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleI 120
    ||||| ..... ||||| ||||| ||||| |||||
210 AGCCACAACCTGATCTCAACCTGCCCAACACAGCTGCTTCATATCC 259
120 euAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136

```

```

|||||
260 TGGCCACCCGCGGTACGAGTGTGGCTGGGAAACGCGGGGCAACACC 309
137 TTPSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
|||||
310 TGGGCGCAGAGAGATCTGACTACTGCGCCGACCTCCGCGAATCTCGGGC 359
153 aPheSerTyrAspGluMetAlaArgPheAspLeuProAlaValIleAsp 170
|||||
360 TTTCAGCTTTGACGAGATGCGTAAATATGACCTCCCGCCACCATGTGACT 409
170 heIleLeuGlnLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSer 186
|||||
410 TCATCTTGAAGAAACGGACAGGACAACTACCTACCTGCGCCCAATTC 459
187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
|||||
460 CAGGGCACCACCATGTGTTTCATCGCTTTCCACCAATCCCAAGCTGGC 509
203 acGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysH 220
|||||
510 GAAACGATCAAAACCTCTATGCAATTAGCTCCCGTGCACACGTAAGT 559
220 IsAlaLysSerProGlyThrLysPheLeuLeuLeuProAspMetMetIle 236
|||||
560 ACACCGAAACCTGTAAACAACTCATGCTGCTCCCTCGCTCTCTTC 609
237 LysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThrArgPheLeuAr 253
|||||
610 AACCTTATATTGGAAACAAATATTC...TACCCACACCACTCTCTTGA 656
253 gGln...LeuValIleTyrLeuGlyGlnValIleLeuAspGlnIleC 269
|||||
657 TCATATTCTCGCCACCGAGGTATGCTCCGCGAGAGGTGATCTCTCT 706
269 ySerSerIleMetLeuLeuLeuGlyGlyPheAsnThrAsnMetAsn 285
|||||
707 GCGACACCGCCTGTTATCATTTGATTTGACACTATGAACCTTGAAAC 756
286 MetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVa 302
|||||
757 ATAGTCGCTGTGATGTATCTGTCACATTAATCCACAGAGAACATCGGT 806
302 IGIAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyLeuLeuArg 319
|||||
807 TCGAACGCTGCTCCACTGCTCCAGGCTGTAACTGTGGAAAGTTCAG 856
319 IapHeAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 335
|||||
857 CTTTGACTGGGGAGCCCACTTCAGAACATGATGCACTATCTCAGAGC 906
336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMetTr 352
|||||
907 ATGCCCTCCCTACTACAACTGACAGACATGCAATGCCAATGCCAATG 956
352 rHnGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuL 369
|||||
957 GAACGCTGGCAACGACTGCTGCGCAGCCCTCAGATGTGACTTTTGC 1006
369 euSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAla 385
|||||
1007 TTTCAGAGCTCCCAATCTCATTTACCAAGGAAGATCTCTCTTCAAT 1056
386 HisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnG 402
|||||
1057 CACTTGGAGCTTATCTGGGCGCATGATGCCCTCAAGCGTTTACATGA 1106
402 uIleIleHisLeuMet 407
|||||
1107 AATGTTCATGATG 1122
seq_name: gp_pat:AR039023

```

```

seq_documentation_block:
LOCUS AR039023 1140 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 4 from patent US 5807726.
ACCESSION AR039023
VERSION AR039023.1 GI:5958386
KEYWORDS
SOURCE
ORIGIN
REFERENCE
AUTHORS 1 (bases 1 to 1140)
TITLE Blanchard,C., Benicourt,C. and Julien,J.
Nucleic acids encoding dog gastric lipase and their use for the
production of polypeptides
JOURNAL Patent: US 5807726-A 4 15-SEP-1998;
FEATURES
source Location/Qualifiers
BASE COUNT 297 a 308 c 250 g 285 t
ORIGIN
alignment_scores:
Quality: 1141.50 Length: 372
Ratio: 3.624 Gaps: 3
Percent Similarity: 84.677 Percent Identity: 56.452
alignment_block:
US-09-578-063-47 x AR039023
Align seg 1/1 to: AR039023 from: 1 to: 1140
37 ProThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIleI 53
|||||
22 CCCACA.....AACCTGAAGTACCACGAAATATGATCAATGATGAT 62
53 eGlnHisGlnGlyTyrProCysGlnGluIuTrpGluValAlaThrGln 70
|||||
63 CACCTACTGGGGATCCCAAGCTGAGAAATATGAAGTTGTGACCAGAGAG 112
70 LysTrpIleLeuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
|||||
113 GTTATATCTTGGGATCGACAGATTCCTTATGGGAGGAAATTCAGAG 162
87 LysThrGlySerArgProValValLeuLeuGlnHisGlyLeuValGly 103
|||||
163 AATATAGCGCGGAGACCTTGTGATTTTTCACACAGGTTTGTCTGCATC 212
103 yAlaSerAsnTrpIleSerAsnLeuProAsnSerLeuGlyPheIleL 120
|||||
213 AGCCACAACTGGATCTCCAACTGCGCAACACACAGCCTGGCCTTATTC 262
120 euAlaAspAlaGlyPheAspValTrpMetGlyAsnSerArgLysAsnAla 136
|||||
263 TGGCGCAGCCCGGGTACGACGTGTGGCTGGGGAACAGAGGGGCAACACC 312
137 TTPSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
|||||
313 TGGGCGCAGAGAGATCTGACTACTGCGCCGACCTCCGTGAATTTCTGG 362
153 aPheSerTyrAspGluMetAlaArgPheAspLeuProAlaValIleAsp 170
|||||
363 TTTCAGCTTTGACGAGATGCGTAAATATGACCTCCCGCCACCATGTGACT 412
170 heIleLeuGlnLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSer 186
|||||
413 TCATCTTGAAGAAACGGACAGGACAACTACCTACCTGCGCCCAATTC 462
187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
|||||
463 CAGGGCACCACCATGTGTTTCATCGCTTTCCACCAATCCCAAGCTGGC 512
203 acGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysH 220
|||||

```

```
513 GAAACGATCAAAACCTTATGCTATAGTCCCGTTGCCACCGTAGT 562
220 IeAlaIysSerProGlyThrLysPheLeuLeuLeuProAspMetMetIle 236
      ::::::::::: ::::::::::: ::::::::::: :::::::::::
563 ACACCGAAACCGTGTAAACAAATCATGCTGCTCCCTTGCTCCTCTTC 612
237 LysGlyLeuPheGlyLysGlyPheLeuTyGlnThrArgPheLeuAr 253
      ::::::::::: ::::::::::: ::::::::::: :::::::::::
613 AAGCTATATTTGGAAACAAATATTC...TACCACACCCACTTCTTGA 659
253 gGln...LeuValIleTyLeuCysGlyGlnValIleLeuAspGlnIleC 269
      ::::::::::: ::::::::::: ::::::::::: :::::::::::
660 TCAATTTCTGCCACCGAGGTATGCTCCCGAGACGGTGGATCTCCT 709
269 ySerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnMetAsn 285
      ::::::::::: ::::::::::: ::::::::::: :::::::::::
710 GCAGCAACGCGCGTTTATCATTTGTGGATTGTGACACTATGAAC 759
286 MetSerArgAlaSerValTyAlaIaIaIaIaIaIaIaIaIaIaIaIa 302
      ::::::::::: ::::::::::: ::::::::::: :::::::::::
760 ATGAGTGGCTGGATGTATCTGTCAATATCCAGACAGAAATCGGT 809
302 IGIaSnIleLeuIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 319
      ::::::::::: ::::::::::: ::::::::::: :::::::::::
810 TCAGAACGTGCTCCACTGCTCCAGGCTGTAAAGTGTGGAGATCCAG 859
319 IaPheAspTyrGlySerGlyThrLysAsnLeuGlnLysCysAsnGlnPro 335
      ::::::::::: ::::::::::: ::::::::::: :::::::::::
860 CTTTGTACTGGGAGACCCGAGTTCAGAACATGATGCATCATCAGAGC 909
336 ThrProValArgTyArgValaIaArgAspMetThrValProThrAlaMetTr 352
      ::::::::::: ::::::::::: ::::::::::: :::::::::::
910 ATGCTCCCTACTACAACTGACACAGACATGATGCGCAATCGCAGTGTG 959
352 pThrGlyGlyGlnAspTyrLeuSerAsnProGluAspValIysMetLeuL 369
      ::::::::::: ::::::::::: ::::::::::: :::::::::::
960 GAACGGTGGCAAGACTGTGCTGCCGACCTCAGAGATGTGACCTTTGC 1009
369 euSerGluValThrAsnLeuIleTyThrHisLysAsnIleProGluTyrPa 385
      ::::::::::: ::::::::::: ::::::::::: :::::::::::
1010 TTCCCAAGCTCCCAATCTCATTTACCAAGAAAGATTCTCTTACAAT 1059
386 HisValAspPheIleTyrGlyLeuAspAlaProHisArgMetTyArgnG 402
      ::::::::::: ::::::::::: ::::::::::: :::::::::::
1060 CACTTGACTTTATCTGGGCAATGATGCCCTCAAGCGGTTACAATGA 1109
402 uIleIleHisLeuMet 407
      ::::::::::: :::::::::::
1110 AATTGTTTCCATGATG 1125
```



```

1 MetLeuGlnThrLeuSerArgGlnTrpIleValSerHisArgMetGlu 17
1 ATGTGGAAACCTGTGTCAAGACAGTGGATGTCTCACACAGATGGAAAT 50
17 tTrrLeuLeuIleLeuValAlaIleTyrMetPheGlnArgAsnValAsnSer 34
51 GTGGCTTCTGATCTGTGGCGGTATATGTCTCAGAGAAATGGAATTTAG 100
34 aHisMetProThrLysAlaValAlaValAspProGluAlaPheMetAsnIleSer 50
101 TACATATGCCAAGCTAAAGCTGTGGACCCAGAACTTCATGAAATTTAGT 150
51 GluIleIleGlnHisGlnGlyTyrProCysGlnGluIleTyrGluValAla 67
151 GAAATCATCCAAACATCAAGAGCTATCCCTGTGAGGAATATGAAGTGGCAAC 200
67 rGluAspGlyTyrIleLeuSerValAsnArgIleProArgGlyLeuValG 84
201 TGAAGATGGTATATCTTTCTGTAAACAGGATTCCTGAGGCTTAGTGC 250
84 InProLysLysThrGlySerArgProValValIleLeuGlnHisGlyLeu 100
251 AACCTAAAGACAGAGCTTCAGAGCCTGTGGTGTACTGCAGCATGGCTA 300
101 ValGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuG 117
301 GTTGAGGTGGTGAAGCACTGATTTCCAACTGCCCAACAATATAGCTTGG 350
117 yPheIleLeuAlaAspAlaGlyPheAspValIleTyrMetGlyAsnSerArg 134
351 CTTCATCTGGCAGATGCTGTTTGCAGCTGTGAGTGGGAAACACAGAGG 400
134 LyAsnAlaIleTrpSerArgLysHisLysThrLeuSerIleAspGlnAspG 150
401 GAAAGCGCTGGCTCGAAACACAGACACTCTCATAGACCACAGATAG 450
151 PheTrpAlaPheSerTyrAspGlnMetAlaArgPheAspLeuProAlaVal 167
451 TTCGGGCTTTCAGTATGATGAGATGGCTAGTGTGACTTCCGCAAGT 500
167 IleAsnPheIleLeuGlnLysThrGlyGlnGlyLysIleTyrValG 184
501 GATAAACTTATTTTGCAGAAACCGGCGCAGAAAGATCTATATATGCG 550
184 LysTyrSerGlnGlyThrThrMetGlyPheIleAlaPheSerThrMetPro 200
551 GCTATTCACAGGGCACACCAATGGCTTATTTGCAATTTCCACCATGGCCA 600
201 GluLeuAlaGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThr 217
601 GAGCTGGCTCAGAAATCAAAATGATTTTGTCTTAGCACCATTAGCCAC 650
217 rValLysHisAlaLysSerProGlyTyrLysPheLeuLeuLeuProAspM 234
651 TGTTAAGCATCAAAAACCCCGGACCAAAATTTTGTGCTGCCAGATA 700
234 eTMetIleLysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThrArg 250
701 TGATGATCAAGGAGTGTGTGGCAAAAAGAAATTTCTATACAGCCAGCA 750
251 PheLeuArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspG 267
751 TTTTCACACAACTGTATATTACCTTGTGGCCAGGTGATTCCTTGATCA 800
267 nIleCysSerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnAsnM 284
801 GATTGTGTATATATCATGTTACTTCTGTGGATTCACACCAACATATA 850
284 eTanMetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThr 300
851 TGAACATGAGCCGAGCAAGTGTATATGCTGCCACACACTTCTTGTGAACA 900

```

```

301 SerValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGlu 317
901 TCTGTGCAAAATATCTACACTGAGCCAGCAGAGTAATCTGTGTAAGT 950
317 uArgAlaPheAspTrpGlySerGlnThrLysAsnLeuLysCysAsnG 334
951 CCGGCAATTTGACTGGGGAGTGAAGCCAAAATCTGGAAAAATGCCAATC 1000
334 InProThrProValArgTyrArgValArgAspMetThrValProThrAla 350
1001 AGCCACTCTCTTAAGGTACAGAGTCAGAGATATACCGTCCCTACACA 1050
351 MetTrpThrGlyGlnAspTrpLeuSerAsnProGluAspValLysMe 367
1051 ATGTGACAGAGAGGTACAGAGCTGGCTTCAATCCAGAAAGCTGAAAT 1100
367 tLeuLeuSerGluValThrAsnLeuIleTyrHisLysAsnIleProGlu 384
1101 GCTGCTCTGAGGTGACCAACTCATCTACATTAAGAATATTCCTGAAT 1150
384 rPalAlaHisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyr 400
1151 GGGCTCACGTGATTTCACTGGGGTTGGATGCTCTCACCGTATGTAC 1200
401 AsnGluIleIleHisLeuMetGlnGlnGluIleThrAsnLeuSerGln 417
1201 AATGAATCATTCATCTGATGAGCAGAGAGAGACCAACTTTCACAGG 1250
417 yArgCysGluAlaValLeu 423
1251 ACGGTGTGAGGCCGTATATG 1269

```

seq_name: /SID58/gcgdata/geneseq/geneseqn/NA2001.DAT.AAF45131

seq_documentation_block:

ID AAF45131 standard; cDNA; 2044 BP.

XX AAF45131;

XX 30-MAR-2001 (first entry)

XX Human TANGO 294 cDNA.

XX Gene therapy: TANGO protein; INTERCEPT protein; neurological disorder;

KM central nervous system; focal brain disorder; bipolar affective disorder;

KM global diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;

KM senile dementia; Huntington's disease; amyotrophic lateral sclerosis;

KM Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;

KM neuropsychiatric; psychoactive substance use; anxiety; ss.

OS Homo sapiens.

XX WO200077239-A2.

XX 21-DEC-2000.

XX 24-MAY-2000; 2000MO-US14858.

XX 14-JUN-1999; 99US-0333159.

XX (MILL-) MILLENNIUM PHARM INC.

XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX WPI; 2001-032313/04.

XX P-PSDB; AAB6065.

XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for

PT screening assays and diagnostic assays and for the treatment of

PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's

PT disease -

XX Claim 1; Fig 6; 359pp; English.

XX The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders;
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.

XX Sequence 2044 BP, 618 A, 401 C, 460 G, 565 T, 0 other;

alignment_scores:
 Quality: 2247.00 Length: 423
 Ratio: 5.312 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 us-09-578-063-47 x AAF45131

Align seg 1/1 to: AAF45131 from: 1 to: 2044

1 MetLeuGluThrLeuSerArgGlnTrpIleValSerHisArgMetGlu 17
 126 ATGTTGGAACCTGTCAAGACAGTGGATGTCTCACACAGATGGAAT 175
 17 TTPRLeuLeuIleuValAlaTyrMetPheGlnArgHisValSer 34
 176 GTGGCTTCTGTGTTGGTGGGTATGATCTCCAGAAATGTGAATCAG 225
 34 AlHisMetProThrLysAlaValAspProGlnIlePheMetAsnIleSer 50
 226 TACATATGCCACTAAAGCTGTGAGCCAGAGCAATTCATATAATATG 275
 51 GluIleIleGlnHisGlnGlyTyrProCysGluGluTyrGluValAla 67
 276 GAATTCATCCACATCAAGGCTATCCCTGTGAGAAATATGAAGTCGAC 325
 67 rGluAspGlyTyrIleLeuSerValAsnArgIleProArgGlyLeuVal 84
 326 TGAAGATGGGTATCTCTCTTAAACAGATTCCTGAGGCTAGTGC 375
 84 InProLysThrGlySerArgProValIleLeuGlnHisGlyLeu 100
 376 AACCTAAGAAAGACAGGTCCAGGCTGTGGTGTACTGACAGATGGCTTA 425
 101 ValGlyIleAlaSerAsnTrpIleSerAsnLeuProAsnSerLeuG 117
 426 GTTGAGAGGTGTAGCAACTGGATTCCAACTGCCCAACATATACCTGG 475
 117 yPheIleLeuAlaAspAlaGlyPheAspValTyrMetGlyAsnSerArg 134
 476 CTTCATCTCTGACAGATCTGTGTTTGAAGTGTGATGGGAAACAGCAGG 525
 134 LysAlaAlaTyrPserArgLysHisLysThrLeuSerIleAspGlu 150
 526 GAAGAGGCTGTCTGAAACACAGACACTCTCATATGACCAAGATGAG 575
 151 PheTrpAlaPheSerTyrAspGluMetAlaArgPheAspLeuProAla 167
 576 TTCTGGGCTTTCAGTATGATGAGATGGCTATGCTTCTGCTGAGT 625
 167 ILeAsnPheIleLeuGlnLysThrGlyGlnGluLysIleTyrTyrVal 184
 626 GATTAACCTTATTTTGCAGAAAGGCGCAAGAAAGATCTATATATGCG 675
 184 LyrSerGlnGlyThrThMetGlyPheIleAlaPheSerThMetPro 200

676 GCTATTCACAGGCGACCATGGGCTTATTTGCATTTCCACCATGCA 725
 201 GluLeuAlaGlnLysIleLysMetTyrPheAlaLeuAlaProIleAla 217
 726 GAGGTGGCTCGAAGAAATCAAAATGATTTTCTTATGACCAACATAGCCAC 775
 217 rValLysHisAlaLysSerProGlyThrLysPheLeuLeuProAsp 234
 776 TGTTAAGCATCAAAAGCCCGGAGCCAAATTTTGTGTGGCCAGATA 825
 234 eMetIleLysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThrArg 250
 826 TGATGATCAAGCATTTGTTGGCAAAAAGATTTCTGTATACACACCA 875
 251 PheLeuArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAsp 267
 876 TTTCACAGCAACTGTGTTATTTACCTTTGTGGCAGGTGATTTGTATCA 925
 267 nileCysSerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnAsn 284
 926 GATTTGTAAGTATATCATGTTACTCTGGGTGATTCACACCAACATA 975
 284 eLAsnMetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThr 300
 976 TGAACATGAGCCGAGCAAGTATATGCTGCCACACTCTGTGGAGACA 1025
 301 SerValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGlu 317
 1026 TCTGTGCAAAATATTTCTACCTGAGGACGAGCAGATGTAATCTGGTAACT 1075
 317 uArgAlaPheAspTrpGlySerGluThrLysAsnLeuGlnLysCysAsn 334
 1076 CCGGGCAATTTGACTGGGGAGTGAACCAAAATCTGAAAAAATGCAATC 1125
 334 InProThrProValArgTyrArgValArgAspMetThrValProThrAla 350
 1126 AGCCACTCTCTGTAAGTACAGAGTACAGATATAGCGTCCCTACACA 1175
 351 MetTrpThrGlyGlnAspTrpLeuSerAsnProGluAspValLys 367
 1176 ATGTGACAGAGAGTACAGACTGGCTTCAATCAGAAAGACGTAAAT 1225
 367 tLeuLeuSerGluValThrAsnLeuIleTyrHisLysAsnIleProGlu 384
 1226 GCTGCTCTGAGGTGACCAACCTCATCTACCAAGAAATTTCTCTGAT 1275
 384 rPAlaHisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyr 400
 1276 GGGCTACGTGATTTCTATCTGGGGTTGGATGCTCTCACCCGATATGAC 1325
 401 AsnGluIleIleHisLeuMetGlnGlnGluLysThrAsnLeuSerGln 417
 1326 AATGAATCATCATCTGATGACAGAGAGAGACCAACCTTTCCAGAG 1375
 417 yArgCysGlnAlaValLeu 423
 1376 ACGGTGTGAGGCCGTATTG 1394
 seq_name: /SIDSR/gcgdata/geneseq/geneseqn/NA2001.DAT:AAH02906
 seq_documentation_block:
 ID AAH02906 standard; DNA; 2493 BP.
 AC AAH02906;
 XX
 AC
 DT 15-JUN-2001 (first entry)
 XX
 DE Human shear stress-response coding sequence SEQ ID NO: 65.
 XX
 KW Human; shear stress-response protein; vascular disease;
 KM arteriosclerosis; ds.
 XX
 OS Homo sapiens.

XX K0200125427-A1.
XX
XX 12-APR-2001.
XX
XX 02-OCT-2000; 2000MO-JP06840.
XX
XX 01-OCT-1999; 99UP-0280976.
XX
XX (KYOWA) KYOWA HAKKO KOGYO KK.
XX (NOJI/) NOJIMA H.
XX
XX Nojima H., Yoshisue H., Odayashi M., Ota T., Kawabata A., Sakurada K.,
XX Kuga T., Sekine S., Nakamura Y., Sugano S;
XX
XX WPI; 2001-266308/27.
XX
XX P-PSDB; AAB90783.
XX
XX DNA sequences, proteins encoded by them and antibodies against them
XX useful in diagnosis and treatment of vascular disease caused by
XX arteriosclerosis -
XX
XX
XX Claim 20; Page 398-402; 678pp; Japanese.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human shear stress response proteins. These are useful in the
XX diagnosis, treatment and screening of vascular diseases caused by
XX arteriosclerosis, including heart failure, post-PTCA restenosis and
XX hypertension.
XX
XX
XX Sequence 2493 BP; 692 A; 484 C; 516 G; 801 T; 0 other;

alignment_scores:

Quality: 1306.00	Length: 410
Ratio: 3.841	Gaps: 2
Percent Similarity: 82.927	Percent Identity: 60.732

alignment_block:

US-09-578-063-47 x AAH02906 ..

Align seg 1/1 to: AAH02906 from: 1 to: 2493

```
2 LeuGIuTrIleuSerArgInTrIleValSerHisArgMetGluMetr 18
8 CTCGAGACAGCGCGCCGAG.....GACAGCTCCAGATGAATGCG 51
18 pleuLeuIleuValAlaIlyrMetPheGlnArgAsnValAsnSerValH 35
52 GTTCTGGGGTGGTGGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGAG 101
35 IsMetProThrIlys.....AlaValAspProGluAlaPheMetAsnIle 49
102 GGTCTGGAGGAGAACTGACAGCTGTGATCTCTGAACAACAATGATG 151
50 SerGIuIleIleGlnHisGlnIleIlyrProCysGlnIlyrGluValAl 66
152 AGTGAATTAATCTTACTGTGGGATTCCTAGTGAAGATACCTAGTGA 201
66 aThrIuAspGlyIlyrIleleuSerValAsnArgIleProArgGlyLeuV 83
202 GACAGAAATGATATATCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 251
83 aGIuProIlySlyThrGIySerArgProValIleuLeuGlnHisGly 99
232 AGAACCATTTCTGACAAAGTCCCAACCAAGTGTCTCTCTGCAACATG 301
100 LeuValGIyGIyAlaSerAsnTrIleSerAsnLeuProAsnAsnSerIe 116
302 TTGCTGGCAGATTTCTAGTAACTGCGTCAAACTTGCACACAGACGCT 351
116 uGIyPheIleIleuAlaAspAlaGlyPheAspAlaTrIlyMetGlyAsnSe 133
```

```
352 GGCCTCATTTCTGTGATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 401
133 rGlyAsnAlaTrpSerArgIlySlySlyIleuSerIleAspGlnAsp 149
402 GAGGAATATACCTGTCTGCGAAGATGAAGACACTCTCAGTTCTTCAGAT 451
150 GIuPheTrpAlaPheSerTrpAspGluMetAlaArgPheAspLeuProAl 166
452 GAATTCGTGGCTTTCAGTATGATGATGATGATGATGATGATGATGATGAT 501
166 aValIleAsnPheIleuGlnIlySlyThrGlyGlnIlySlyIleTrpYrV 183
502 TTCCATTACTTCTGATGATAAATGCTGCGCAACACAGTGTATTG 551
183 aGIlyrSerGlnIlyrThrMetGlyPheIleAlaPheSerThrMet 199
552 TGGTCATTTCTCAAGCAGCAGCATATAGTATTATAGCATTTTTCACAGAT 601
200 ProGIuLeuAlaGlnIlySlyMetIlyrPheAlaLeuAlaProIleAl 216
602 CCTGAGCTGGCTAAAGATTAATGTTTGGCTGGCTGGCTGGCTGGCTGG 651
216 aThrValIlySlyHisAlaIlySerProGIyTrIlyPheIleuLeuProA 233
652 TTCCTGCCCTTCTGTACTAGCCCTATGCGCAATTAAGACGATTACCG 701
233 sPmetMetIleSlyGIySlyPheGIySlyGlnPheLeuTrIlyr 249
702 ATCATCTCATTAAGACCTTATTGGAGACAAAGATTTCTTCCACAGAT 751
250 ArgPheLeuArgGlnLeuValIleTrIlyLeuCysGlnValIleLeuAs 266
752 GCGTTTGAAGTGGCTGGGTACCCAGCTTTCACATGATCATGATGATGA 801
266 pGIuIleCysSerAsnIleMetLeuLeuGlyIlyPheAsnTrIlyrAsn 283
802 GGAGCTGTGGAAATCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 851
283 sMetAsnMetSerArgAlaSerValTrpAlaAlaHisTrIlyLeuAlaGly 299
852 ATTTAAATATGCTAGAGTGTATATACACACATTTCTCTCTGCTGGA 901
300 ThrSerValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGly 316
902 ACTTCTGTGCAAAACATGATGTACATGAGCCAGCTGTTAAATTCGAA 951
316 uLeuArgAlaPheAspTrpGIySerGIuTrIlyAsnLeuGluIlyScySA 333
952 GTTTCAGCCCTTGTACTGGGGAAGCAGTGCACAAAGATTAATTTTCATT 1001
333 sngIProThrProValArgTrpArgValArgAspMetThrValProThr 349
1002 ACCAAGATTATCTCCACATACATGTAAGAGATATGTTTGGCGACT 1051
350 AlaMetTrpTrpGIyGlnAspTrpLeuSerAsnProGluAspValIly 366
1052 GCAGCTGTGAGCGGGGTACAGACAGCTGCTGTGAGATGCTACAGAGCAA 1101
366 sMetLeuLeuSerGIuValIleAsnLeuIleTrIlyHisAsnIleProG 383
1102 TATCTTACTGACTGACATCACCAACTGTGTCTCATGAGACATTCGG 1151
383 IuTrpAlaHisValAspPheIleTrpGIyLeuAspAlaProHisArgMet 399
1152 AATGGAGACATTTCTCATTTTGGGCTGTGATGCCCTTGGAGGCTT 1201
400 TyrAsnGIuIleIleHisLeuMetGlnGln 409
1202 TATAATAAATTAATTAATTAATGAGAA 1231
seq_name: /SIDSB/gcgdata/geneseq/geneseq/NA1993.DAT:AA042310
seq_documentation_block:
```



```

360 SerAspProGluAspValIleuMetLeuSerGluValThrAsnLeu1 376
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
1089 GGTGACCCCAAGATGTTGACCTTGTGCTCCAAACCTTCAATCTCAT 1138
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
376 eTyHisLysAsnIleProGluTrpAlaHisValAspHeileTrpGlyL 393
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
1139 TTACCACAAGAGATTTCCATATAACACTTGATTTATCTGGGCAA 1188
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
393 euAspAlaProHisArgMetLysGluIleIleHisLeuMetGlnGln 409
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
1189 TGAATGCCCTCCACAGAGTTTACATGAATATTTCTATGATGGCAAAA 1238
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
410 GluGlu 411
      ::::::::::::::
1239 GATATA 1244

seq_name: /SIDS8/gcgdata/geneseq/geneseq/NA1985.DAT: AAN50385
seq_documentation_block:
ID   AAN50385 standard; DNA; 1336 BP.
XX
AC   AAN50385;
XX
DT   17-JAN-1992 (first entry)
XX
DE   Rat prelingual lipase gene.
XX
KW   Prelingual lipase; enzyme; EC-3.1.1.3; ss.
XX
OS   Rattus rattus.
XX
FH   Key          Location/Qualifiers
FT   CDS          7..1194
FT               /*tag= a
XX
FT   GB2142337-A.
XX
PD   16-JAN-1985.
XX
PF   29-JUN-1984; 84GB-0016581.
XX
PR   01-JUL-1983; 83GB-0017989.
PR   05-SEP-1983; 83GB-0023759.
XX
PA   (CELL-) CELLTECH LTD.
XX
PI   Carey NH, Williamson R;
XX
DR   WPI: 1985-014450/03.
DR   P-PSDB; AAP50322.
XX
PT   New lingual lipase protein for treatment of lipase deficiency - also
XX   new pre-lingual lipase protein and related products
XX
PS   Disclosure; Fig 5; 15pp; English.
XX
CC   This sequence may be expressed in a transformant host organism using
CC   a vector plasmid in order to produce a prelingual lipase protein.
CC   The expressed protein may be used for the treatment of lipase
CC   deficiency.
XX
SO   Sequence 1336 BP; 409 A; 285 C; 263 G; 379 T; 0 other;

alignment_scores:
      Quality: 1158.50      Length: 399
      Ratio: 3.543          Gaps: 1
Percent Similarity: 81.955 Percent Identity: 54.386

alignment_block:
us-09-578-063-47 x AAN50385 ..
Align seg 1/1 to: AAN50385 from: 1 to: 1336

```

```

16 GluMetTrpLeuLeuIleLeuValAlaTrpMetPheGln.. ArgAsnVal 31
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
4   AAGATGTGCTGTATTATATAACCAAGTGATCAACATTGGAGTGC 53
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
31 LAsnSerValHisMetProThrLysAlaValAspProGluAlaPheMet 48
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
54 ACATGGCCATTATGGAAAACTGGTCCTGGAAAACTTGAGCAAAATATGA 103
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
48 snIleSerGluIleIleGlnHisGlnGlyTrpProCysGluIuTrpGlu 64
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
104 ATATTAGTCAGATGATTAATTACTGCGGATATCCATGTCAGAAATATGA 153
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
65 ValAlaThrGluAspGlyTrpIleLeuSerValAsnArgIleProArgL 81
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
154 GTTGTACTAGAGATGGCTACATTTCTGGGGCTCTACAGATTCCTCATG 203
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
81 YLeuValGlnProLysLysThrGlySerArgProValIleLeuGlnH 98
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
204 GAAGAAATATTCTGAAAAATATAGCGACAGACCTGTGTATTGCGAC 253
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
98 ISGlyLeuValGlyLysAlaSerAsnTrpIleSerAsnLeuProAsn 114
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
254 ATGTTTGTATGCGATCGACAGCAAACTGATGCAATCTACCAACAAC 303
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
115 SerLeuGlyPheIleLeuAlaAspAlaGlyPheAspValTrpMetGly 131
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
304 AGCCCTGCGCTTATGTCAGACACGCTGCTATGATGTGTGCTGGGGA 353
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
131 nSerArgGlyAsnAlaTrpSerArgLysHisLysThrLeuSerIleAsp 148
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
354 CAGTCGAGAAATACATGTCAGAGAAATATATACACTACCAAGACT 403
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
404 CAGTTGAATTCGGGCTTCACCTTGATGAATGCTAATATAGCCT 453
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
165 ProAlaValIleAsnPheIleLeuGlnLysThrGlyGlnGlyLysIle 181
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
454 CCGGCCACAAATAACTTGTACAGAAACTGACAGAAAGAAATATACA 503
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
181 rTyrValGlyTyrSerGlnGlyThrThreGlyPheIleAlaPheSer 198
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
504 CATATTGCTCATTTCTACGCGCACCATATTGTTTCATTGCTTTTCA 553
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
198 hMetProGluLeuAlaGlnLysIleLysMetLysPheAlaLeuAlaPro 214
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
554 CCAATCCTACACTGCGCAAAAAATCAAGACGTTTATGCAATTACTCCA 603
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
215 IleAlaThrValLysHisAlaLysSerProGlyThrLysPheLeuLeu 231
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
604 GTTGTACCGTGAAGTATACACAAAGCTTGAAAAAGATTTTCATTAT 653
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
231 uProAspMetIleLysGlyLeuPheGlyLysLysGluPheLeuTrp 248
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
654 TCTTACATTCTTTCACAGCTTATGTCGCAAGAAATGTTCCGCCCC 703
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
248 InThrArgPheLeuArgLysLeuValIleTyLeuCysGlyGlnValIle 264
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
704 ACACCTACTTTGATGACCTTCTGTGTACCGAAGTGTGCTACGGAGGT 753
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
265 LeuAspGlnIleCysSerAsnIleMetLeuLeuGlyGlyPheAsnTh 281
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
754 CTAGATCTTCTCTGCGACACACTTATTCATCTCTGCTGTGATTCACA 803
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
281 rAsnAsnMetAsnMetSerArgAlaSerValTrpAlaHisThrLeuA 298
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
804 GAAAAACTTAATGAGTGTGTTTGTATGTATAGGCAATATCCAG 853
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
298 IaglyThrSerValGlnAsnIleLeuHisLysTrpSerGlnAlaValAsn 314
      :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
854 CAGGACATCTGTTCACAGACTTTCACACTGGCGACACTGTGTAGACT 903

```

[illegible]

657 TCAATTTCGCCACACGAGATGTCGCCGAGACGGTGCATCTCT 706
 269 ySserAsnIleMetLeuLeuGlyGlyPheAsnThrAsnMetAsn 285
 707 GCAGACAGCCCTGTTATCATTTGTGATTTGACACATGTAACCTGGAAC 756
 286 MetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVal 302
 757 ATGAGTCCTGCTGATGTATCTGTACATATATCCAGACGAGAACTCCGT 806
 302 IGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArg 319
 807 TCAGACAGTGCCTCCACTGCTGCCAGCGCTGTAACTGCGAAGTCCACAG 856
 319 IapheAspTrpGlySerGluThrLysAsnLeuGlyLysCysAsnGlnPro 335
 857 CTTTGGATGGGGAGGCCAGCTTCAGAACATGATGCATATCATCATCAGAC 906
 336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMetTr 352
 907 ATGCCTCCCTACTACACCTGACAGACATGATGCGCAATCGCATGTG 956
 352 PThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeuL 369
 957 GACGGTGGCAACGACTGCTGGCCGACCTCCACATGTTGACCTTTGC 1006
 369 euserGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrrpAla 385
 1007 TTTCACAGCTCCCAATCTCATTTACACAGGAAGATCTCCCTTACAAAT 1056
 386 HisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnG 402
 1057 CACTTGACCTTATCTGGCCATGATGATCCCTCAAGCGTTTACAATGA 1106
 402 uIleIleHisLeuMet 407
 1107 AATGTTTCCATGATG 1122

seq_name: /SID8/9c9data/geneseq/geneseqn/NA1996.DAT.AAT58915

seq_documentation_block:
 ID AAT58915 standard; cDNA to mRNA; 1528 BP.
 AC AAT58915;
 DT 19-AUG-1997 (first entry)
 DE Dog gastric lipase coding sequence.
 XX
 KW Dog; duodenal; gastric; lipase; transgenic; plant; recombinant; extract;
 food; absorption; fat; pancreatic; cystic fibrosis; exocrine; dairy;
 hydrolysis; trans-esterification; substrate; enzyme; biofuel; ss.
 OS
 XX
 FH Canis familiaris.
 FT key
 FT Location/Qualifiers
 CDS 1..1140
 /tag- a
 /product- gastric lipase

W09633277-A2.
 XX
 PD 24-OCT-1996.
 XX
 PF 19-APR-1996; 96MO-FR00606.
 XX
 PR 20-APR-1995; 95FR-0004754.
 XX
 PA (BIOC-) BIOCEM SA.
 PA (LJOU) INST RECH JOUVEINAL.
 PI Baudino S, Benicourt C, Cudrey C, Gruber V, Lenee P;
 Merot B;

XX
 DR WPI: 1996-485783/48.
 DR P-PSDB; AAM09382.
 XX
 PT Recombinant human or canine pre-duodenal lipase prodn. in transgenic
 PT plants - useful for facilitating absorption of fat, as bio-catalysts
 PT and for prodn. of bio-fuel
 XX
 PS Claim 2; Fig 1; 130pp; French.
 PS
 XX
 CC This is the nucleotide sequence encoding the dog pre-duodenal (i.e.
 CC gastric) lipase enzyme. The sequence can be used to generate transgenic
 CC plants producing recombinant lipase in an enzymatically active form.
 CC Alternatively the first 12 bases of this sequence (encoding amino acids
 CC 1-4 of the protein encoded by this sequence) or bases 1-162 of this
 CC sequence (encoding amino acids 1-54) can be deleted to form the
 CC derivatives designated delta-4 or delta-54 respectively. Plants, or
 CC their extracts, expressing the lipases or the truncated derivatives, can
 CC be used: (a) as pharmaceuticals or food to facilitate absorption of fat,
 CC either in healthy subjects or in patients with inadequate levels of
 CC gastric/pancreatic lipase e.g. those with cystic fibrosis or exocrine
 CC pancreatic insufficiency, in the elderly or in patients undergoing medical
 CC treatment which alters fat absorption; (b) for performing industrial or
 CC agricultural reactions, e.g. in processing of fats or in the dairy
 CC industry, for hydrolysis or transesterification reactions, etc., where
 CC the plant material may provide both enzyme and substrate. The transgenic
 CC plants can also be used for biofuel production.
 XX
 SQ Sequence 1528 BP; 392 A; 387 C; 329 G; 420 T; 0 other;

alignment_scores:
 Quality: 1141.50 Length: 372
 Ratio: 3.624 Gaps: 3
 Percent Similarity: 84.677 Percent Identity: 56.452

alignment_block:
 US-09-578-063-47 x AAT58915

Align seg 1/1 to: AAT58915 from: 1 to: 1528

37 ProThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIleI 53
 19 CCCACA.....AACCTGAAGTGACATGATATATACATATAT 59
 53 eGlnHisGlnGlyTyrProCysGlnGluTyrGluValAlaThrGluAsp 70
 60 CACCTACTGGGGATRACCACTGAGAGATATGAACTTGTGACCGAAGACG 109
 70 IYTYrIleLeuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
 110 GTTATATCTTGGATGACAGAAATTCCTTATGGAGAGAAAAATTCAGAG 159
 87 LysThrGlySerArgProValIleLeuLeuGlnHisGlyLeuValGly 103
 160 AATATAGCCGCGAGACCTGTTCATTTTTCACACAGGTTGCTGCATC 209
 103 YAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIle 120
 210 AGCCACAACCTGATCTCCACACCTGCCACACAGCCTGCGCTTCATCC 259
 120 eValAspAlaGlyPheAspValTrrpMetGlyAsnSerArgGlyAsnAla 136
 260 TGGCCGAGCGCGGTACGAGCTGTGGTGGGAGACAGCGGGGCAACACC 309
 137 TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAl 153
 310 TGGCCAGAGGAATCTGTACTACTCGCCGACCTCGTGAATTCGGGC 359
 153 aPheSerTyrAspGluMetAlaArgPheAspLeuProAlaValIleAsn 170
 360 TTTCAGCTTTCACGATGAGTGAATATATGACCTTCCGCCACATGATGACT 409

```

170 heileuGlnlyThrGlyGlnGluLysIleTyrValGlyTyrSer 186
187 GlnGlyThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
460 CAGGGACACACCATGTGTTTCATCGCCTTTCCACCAATCCAGAGTGGC 509
203 aglnlyIleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysH 220
510 GAACCGATCAAAACCTCTATCATTAAGTCCCGTGCACCGGTGAAGT 559
220 IsaLysSerProGlyThrLysPheLeuLeuLeuProAspMetMetIle 236
560 ACACCGAAACCTGTTAAACAACTCATGCTGCCCTTGCTGCTCTTC 609
237 LysGlyLeuPheGlyLysGlyLysGlyLysGlyLysGlyLysGlyLys 253
610 AACGTTATATTGGAAACAAATATTC...TACCCACACACATTTTGA 656
253 ggl...LeuValIleTyrLeuGlyGlyGlyGlyGlyGlyGlyGlyGly 269
657 TCAATTTTCGCCACGAGGATATGCTCCGCGAGAGGAGTGCCTCT 706
269 ySerAsnIleMetLeuLeuGlyGlyGlyPheAsnThrAsnMetAsn 285
707 GCAGCAACGCCCTTATCATTTGGATTGACACTATGACCTTGAAAC 756
286 MetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaLysThrVal 302
757 ATAGTCGCTGGATGATGATCTGCATCAATATCCAGACGAAACATCGGT 806
302 IGlAsnIleLeuHisTyrSerGlnAlaValAsnSerGlyLysLeuArg 319
807 TCAGAACGTCCTCCACTGCTGCCAGGCTGTTAAGTCTGGGAAGTCCAG 856
319 IAPheAspTyrGlySerGlyThrLysAsnLeuGlyLysCysAsnGlnPro 335
857 CTTTGTACTGGGAGAGCCAGTTCAGAACATGATGCATTCATCAAGC 906
336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMet 352
907 ATGCCCTCCCTACTACACCTGACAGACATCATGATGCCAATCGAGTGG 956
352 pThrGlyGlyGlnAspTyrPheLeuSerAsnProGluAspValLysMetLeu 369
957 GAGCGGTGGGAGACACTGCTGCCAGCCCTCAAGATGTTGACCTTTGC 1006
369 euSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTyrPala 385
1007 TTTCCAGGCTCCCAATCTCATTTACCACAGAAAGATTCCTCTTACAAAT 1056
386 HisValAspPheIleTyrGlyLeuAspAlaProHisArgMetTyrAsnG 402
1057 CACTTGACTTATCTGGCCATGATGATGCCCTCAAGCGGTTTACAAATGA 1106
402 uIleIleHisLeuMet 407
1107 AATTGTTTCCATGATG 1122
seq_name: /SID88/gcdata/geneseq/geneseqn/NA1994.DAT:AA068389
seq_documentation_block:
ID AA068389 standard; DNA; 1531 BP.
XX AA068389;
DT 20-FEB-1995 (first entry)
XX
DE Canine gastric lipase (Met-CGL).
XX
KW Canine gastric lipase; CGL; dog; fat; mucoviscidosis; enzyme;
bicoconversion; exocrine pancreatic insufficiency; ss.

```

```

XX OS Canis familiaris.
XX FH Key
XX FT CDS 1..1143
XX FT CDS /*tag= a
XX PN W09413816-A.
XX 23-JUN-1994.
XX 16-DEC-1993; 93MO-FR01260.
XX 16-DEC-1992; 92FR-0015201.
XX PA (LJ00 ) INST RECH JOUVEINAL.
XX PI Benicourt C, Blanchard C, Junien J;
XX DR WPI: 1994-217890/26.
XX DR P-PSDB; AAR56871.
XX PT Recombinant canine gastric lipase and nucleic acid encoding it
XX PT are used for improving absorption of ingested fat, treating
XX PT mucoviscidosis etc. and in enzymatic bio-conversions
XX PS Claim 2; Fig 8; 52pp; French.
XX
The sequence given below is the sequence of figure 8, altered
according to the amendments described on page 2 of the appended
letter.
CC CGL is used to improve absorption of ingested fat, in healthy and
CC sick patients (e.g. having altered levels of gastric lipase); to
CC treat conditions associated with insufficiency (or lack) of lipases,
CC esp. mucoviscidosis or exocrine pancreatic insufficiency and partic.
CC where immobilised, for bicoconversions, e.g. hydrolysis or
CC transacetylation (other mammalian gastric lipases, or derivs.,
CC can be used in this application).
XX
Sequence 1531 BP; 395 A; 386 C; 329 G; 421 T; 0 other:

alignment_scores:
Quality: 1141.50 Length: 372
Ratio: 3.624 Gaps: 3
Percent Similarity: 84.677 Percent Identity: 56.452

alignment_block:
US-09-578-063-47 x AA068389 ..
Align seg 1/1 to: AA068389 from: 1 to: 1531

37 ProThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIleI 53
||||| ..... ||||||| ||||||| .....
22 CCCACA.....AACCTGAAGTACCATCATTAATTAAGTCAGATGAT 62
53 eglnhisGlnGlyTyrProCysGlnGluGlyGlyValAlaLysAspG 70
| : : : ||||||| : : : ||||||| : : :
63 CACCTATGGGGAATACCAGCTGAGGAATATGAAGTGTGACCGAAGACG 112
70 IyTyrIleLeuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
||||| : : : ||||||| : : : ||||||| : : :
113 GTTATATCTTGGGATGACAGAAATTCCTTATGAGGAGAAATTCAGAG 162
87 LysThrGlySerArgProValIleLeuGlnHisGlyLeuValGlyG 103
: : : ||| : : : ||||||| : : : ||||||| : : :
163 AATATAGGCCGGAGACCTGTTCATTTTTCACACAGCTTTGCTCGCATC 212
103 yAlaSerAsnTyrIleSerAsnLeuProAsnAsnSerLeuGlyPheIle 120
: : : ||||||| : : : ||||||| : : : ||||||| : : :
213 AGGCACAAACTGATCTCCAACTGCCCAACACAGCTGCTGCTTCATCC 262
120 euAlaAspAlaGlyPheAspValIleThrMetLysAsnSerArgGlyAsnAla 136

```

```

|||||
263 TGGCCGACCCCGGTACGACGTGTGGCTGGGGAACAGCGGGGCAACACC 312
137 TTPSerArgLysHisLysThrLeuSerIleAspGlnAspGlnPheTrpAl 153
313 TGGGCGAGGAGGAATCTGACTACTGCGCCGACCTCGGTGAATTCGTGGGC 362
153 aPheSerTyrAspGlnMetAlaArgPheAspLeuProAlaValIleAsn 170
363 TTTCAGCTTTGACGAGATGGCTAAATATGACCTCCCGCCACCATGTGACT 412
170 heIleuGlnLysThrGlnGlnLysIleTyrTyrValLysTyrSer 186
413 TCATCTTGAAGAAACGGGACAGACAGACTACCTACTGTGGCCATTC 462
187 GlnGlyThrMetGlyPheIleAlaPheSerThrMetProGlnLeuAl 203
463 CAGGCGACACCATTTGGTTTCATCGCCTTTCCACCAATCCCAAGCTGGC 512
203 aGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaTrpValLysH 220
513 GAAACGGATCAAAACCTTCTATGACATAGCTCCGTCGACACCTGAAGT 562
220 IsAlaLysSerProGlyThrLysPheLeuLeuLeuProAspMetIle 236
563 ACACGGAACCTGTTAAACAACATCATGCTGTCCCTTCGTTCTCTTC 612
237 LysGlyLeuPheGlyLysLysGlnPheLeuTyrGlnThrArgPheLeu 253
613 AAGCTTATATTGGAAACAAATATTC..TACCAACACCATCTTCTTCA 659
253 gGln..LeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleC 269
660 TCAATTTCTCGCCACCGAGGTATGCTCCGCGACAGCGGAGTCTCTCT 709
269 ySerAsnIleMetLeuLeuGlnGlyLysPheAsnThrAsnMetAsn 285
710 GCAGCAAGCCCTGTTATCATTTGTGGTTTGACACTGTGACTTGAAAC 759
286 MetSerArgAlaSerValTyrAlaIleHisThrLeuAlaGlyThrSerVa 302
760 ATGAGTCCCTGGATGTATCTGCACATATCCAGCAGAGAACTCGGT 809
302 IGlAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyLeuLeuArg 319
810 TCAGAACGTGCTCCACTGTCTCCAGGCTGTTAAGTCTGGGAATGCCAAG 859
319 lAPheAspTyrGlySerGlnThrLysAsnLeuGlnLysCysAsnGlnPro 335
860 CTTTGTACTGGGGAAGCCAGTTCAGAACATGATGCACTATCATCAGAGC 909
336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMetTr 352
910 ATGCCTCCCTACTACAACCTGACAGACATGCATGCGCAATCGCAGTGTG 959
352 pThrGlyGlyGlnAspTyrPleuSerAsnProGlnAspValLysMetLeuL 369
960 GAACGGTGCAGACGACTGTGGCCGACCTCACAGATGTTCAGCTTTTGC 1009
369 euSerGlnValThrAsnLeuIleTyrHisLysAsnIleProGlnTrpAla 385
1010 TTTCACACTCCCAATCTCATTTTACCAAGAGATTCCTCTTACAT 1059
386 HisValAspPheIleTyrGlyLeuAspAlaProHisArgMetTyrAsnG 402
1060 CACTTGACCTTATCTGGGCAATGATGCCCTCAAGCGTTTACAATGA 1109
402 uIleIleHisLeuMet 407
1110 AATGTTTCATGATG 1125
seq_name: /SIDS8/gcgdata/geneseq/geneseqn/NA1986.DAT.AAN60685

```

```

seq_documentation_block:
ID      AAN60685 standard; DNA, 1367 BP.
XX
AC      AAN60685;
XX
DE      08-JUN-1991 (first entry)
XX
DE      Sequence encoding pregastric lipase.
XX
KW      Precursor polypeptide; secretion vector; enzyme; ss.
XX
OS      Homo sapiens.
XX
FH      Key
FT      sig_peptide 47..103
FT      mat_peptide 104..1243
FT      /*tag- a
          /*tag- b
XX
WO8603778-A.
XX
PD      03-JUL-1986.
XX
PF      23-DEC-1985; 85WO-GB00599.
XX
PR      21-DEC-1984; 84GB-0032483.
PR      23-DEC-1985; 85WO-GB00599.
PR      01-JAN-1986; 86GB-0019568.
XX
PA      (BREM-) BREWING RES FOUND.
PA      (TUBB/) TUBB R S.
XX
PI      Tubb RS;
XX
DR      WPI: 1986-182910/28.
DR      P-PSDB; AAP60724.
XX
PT      New precursor polypeptide of defined sequence - and corresp. DNA
PT      used to transform hosts for prodn. of the polypeptide
XX
PS      Example; Fig 5; 60pp; English.
XX
CC      The inventors claim a DNA sequence encoding AAP60723 linked to a
CC      promoter upstream and a gene for a polypeptide downstream.
CC      Particular examples are the yeast enzyme AMG, the mammalian enzyme,
CC      gastric lipase and the mammalian lymphokine, interferon-alpha2.
XX
SQ      Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;

alignment_scores:
Quality: 1124.00 Length: 405
Ratio: 3.416 Gaps: 2
Percent Similarity: 81.235 Percent Identity: 52.099

alignment_block:
us-09-578-063-47 x AAN60685 ..

Align seg 1/1 to: AAN60685 from: 1 to: 1367

14 ArgMetGlnMetTyrPleuLeuIleLeuValAlaTyrMetPheGlnArgAs 30
||| ::::::::::::::::::::: ::::| :::::
38 AGTCCAAATGTGGCTGCTTTTAAACATGGCAAGTTGATA..... 79
30 nValAsnSerValHisMetProThrLysAlaVal..... 41
80 .....TCTGTACTGGGCACTACACATGTTTGGAAATATACATC 122
42 .....AspProGlnAlaPheMetAsnIleSerGlnIleGlnHisGln 56
:::|||||: ::::::::::::::::::::: ::::
123 CTGGAAGCCCTGAAGTGTGACTATGAACATTAATGACATATTAATG 172
57 GlyTyrProCysGlnGlnTyrGlnValAlaThrGlnAspGlyTyrIleLe 73

```

173	GGATTCCCAAAATGAAAGATATGAAAGTTGTGACTGAAATGGTTATPATCT	222
173	userValasnaArgIleProArgIleValuGlnProLysIsthArgIys	90
223	TGAAGTCAAATGAGATTCCTTATGGGAAGAAAAATTCAGGGAATACAGCC	272
90	erArgProValValLeuLeuGlnIstGlyLeuValGlyGlyValaSerAsn	106
273	AGAAACCGTTGTGTTTTTGGAGATGGTTTTGCTTGCATCGCACAAAC	322
107	TrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAspAl	123
323	TGGATTTCGAACCTGCGCAACAACAGCCTTCCCTTCATTCTGGCGAGATC	372
123	agIlyPheAspValTrpMetGlyAsnSerArgGlyAsnaIatPserArgL	140
373	TGGTTATATGTGTGGCGGCGCAACAGACAGAGAAACACCGGGCGCAAA	422
140	ysHsIstVstHrLeuSerIleAspGlnAspGluPheTrpAlaPheSerTy	156
423	GAACCTGTACTATTATCCACCAATGCAGTTCAATTCGTGGCTTCAGCTTT	472
157	AspGluMetAlaArgPheAspLeuProAlaValIleAsnPhelIleuG	173
473	GATGAAAGGGTAAATATGACCTTCACACCAACATGCATCATGTTAA	522
173	nlYstHcIyGlnGluLysIleTyrrTyValGlyTySersGlnGlyThr	190
523	GAAACTGACAGAGAAGCAGCTACACTATGTTGGCCATTCACAGGACCA	572
190	hrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLysIle	206
573	CCATGGTTATATGTCCTTTCCACCAATCCCAGCCTGCGCAAAAGATC	622
207	LysMetTyrrPheAlaLeuAlaProIleAlaThrValLysHsIstAlaLysSe	223
623	AAAACCTTCATGCTCATGCTCAGCTCGTTCCACTGGAAGTATACAAAAG	672
223	rProGlyThrIstPheLeuLeuLeuProAspMetMetIleLysGlyLeuP	240
673	CCTTATAAACAACACTAGATTGTTCCCAATCCCTCTTCAAGTTATAT	722
240	heGlyLysLysGluPheLeuTyrglnThrArgPheLeuArgGluLeuVal	256
723	TGGGACAAATATATTACCCACCACTCTTGATCAAAATTCCTTCT	772
257	IleTyrrLeuGysGlyGlnValIleLeuAspGlnIleCysSerAsnIleMe	273
773	ACTGAAGGTGCTCCCGGAGATGCTGATATCTCTTCCACCAAGCTCT	822
273	tLeuLeuLeuGlyGlyPheAsnThrAsnAsnMetAsnMetSerArgIas	290
823	ATTATATATTTGTGGATTGTGACAGTAAACACTTTTAAACGAGATGCTTGG	872
290	erValIlyAlaAlaHisIstPheAlaGlyTrpSerValGlnAsnIleLeu	306
873	ATGTGTATCTATCATATATCCACGAGGAACTTCTGTTCAAAACATGTTC	922
307	HisTrpSerGlnAlaValAsnSerGlyLuleuArgAlaIleAspTrpG	323
923	CATTGGACCCGAGCGTGTAACTCTGGGAAATTCCAACGTTATGACTGGG	972
323	ysSerGlnThrIstAsnLeuGluLysCysAsnGlnProThrProValArgT	340
973	AAGCCCACTTGAATAGATGAGTACATGATACAGTCCCAACCTCCCTACT	1022
340	yrArgValArgAspMetThrValProThrAlaMetTrpThcIyGlyGln	356
1023	ACAATGTACACGCCATGAAATGTACCAATTCGACAGTGTGAACGGTGGCAAG	1072
357	AspTrpLeuSerAsnProGluAspValLysMetLeuLeuSerGluValTh	373

```

1073 GACCGTGGCTGACCAGATGTGGCCCTTTGCCTTCACAAACTGCC 1122
373 rasmnuleletyrnhsllysasnllperoglutpralnlhsvalasphneI 390
|||||.....|||... |||....|||
1123 CAACTATTATTACCAAGAGGAATTCCTTTTACAATCAGTGAGACTTTA 1172
390 letirgylvleuasplaprophisarrghmetYrAsngluilellelshleu 406
|||||.....|||.....:|||||::: ::
1173 TGTGGGCAATGATGATCCCCCTCACAGAAGTTTACATGACATGTTTCATG 1222
407 Metgingingluglu 411
::::::::::
1223 ATATCAGAAAGATAAA 1237

seq_name: /SIDS8/gcgdata/geneseq/geneseqn/AAL1986.DAT.AAN60566
seq_documentation_block:
ID AAN60566 standard; DNA; 1367 BP.
XX .AAN60566;
AC
XX
DT 22-AUG-1991 (first entry)
DE Sequence encoding human pregastric lipase.
XX
KW Cystic fibrosis therapy; enzyme; lipase deficiency; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 47..103 /*tag= a
FT mat_peptide 104..1243
FT /*tag= b
XX
FN WO601532-A.
XX PD 13-MAR-1986.
XX PF 15-AUG-1985; 85MO-G800364.
XX PR 21-AUG-1984; 84GB-0021210.
PR 15-AUG-1985; 85MO-G800364.
PR 01-JAN-1986; 86GB-0008897.
XX PA (CELL-) CELUTECH LTD.
PA (LOWE/) LOWE P A.
PI Lowe PA;
DR WPI: 1986-081634/12.
P-PSDB; AAP60658.
XX New gastric lipase protein, esp. of human origin - for treating
PT lipase deficiency, and DNA sequences coding for it
PS Disclosure; Fig 3; 39pp; English.
XX The inventors claim a pregastric lipase protein and a gene encoding
CC it. Gastric lipase protein is useful for oral administration to
CX treat lipase deficiency, e.g. cystic fibrosis or pancreatitis.
SO Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;

alignment_scores:
Quality: 1124.00 Length: 405
Ratio: 3.416 Gaps: 2
Percent Similarity: 81.235 Percent Identity: 52.099

alignment_block:
US-09-578-063-47 x AAN60566 ..
```


PT plants - useful for facilitating absorption of fat, as bio-catalysts
 PR and for prodn. of bio-fuel

PS Claim 3; Fig 4; 130pp; French.

XX This is the nucleotide sequence encoding the human pre-duodenal (1.e.
 CC gastric) lipase enzyme. The sequence can be used to generate transgenic
 CC plants producing recombinant lipase in an enzymatically active form.
 CC Alternatively bases encoding amino acids 20-23 or 20-73 (1-4 or 1-54 of
 CC the mature protein, respectively) can be deleted to form the derivatives,
 CC designated delta-4 or delta-54 respectively. Plants, or their extracts,
 CC expressing the lipases or the truncated derivatives, can be used:
 CC (a) as pharmaceuticals or food to facilitate absorption of fat, either
 CC in healthy subjects or in patients with inadequate levels of
 CC gastric/pancreatic lipase e.g. those with cystic fibrosis or exocrine
 CC pancreatic insufficiency, in the elderly or in patients undergoing medical
 CC treatment which alters fat adsorption; (b) for performing industrial or
 CC agricultural reactions, e.g. in processing of fats or in the dairy
 CC industry, for hydrolysis or transesterification reactions, etc., where
 CC the plant material may provide both enzyme and substrate. The transgenic
 CC plants can also be used for biofuel production.

XX Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;

alignment_scores:
 Quality: 1124.00 Length: 405
 Ratio: 3.416 Gaps: 2
 Percent Similarity: 81.235 Percent Identity: 52.099

alignment_block:
 US-09-578-063-47 x AAT58916 ..

Align seg 1/1 to: AAT58916 from: 1 to: 1367

14 ArgMetGluMetTrpLeuLeuLeuValAlaTyrMetPheGlnArgAs 30
 38 AGGTCCAAATGTGGCTGCTTTTACAAATGGCAAGTTGATA..... 79
 30 nValAsnSerValHisMetProThrIlyAlaVal..... 41
 80TCTGTACTGGGACCTACACATGGTTGTTGAAATTTACATC 122
 42AspProGluAlaPheMetAsnIleSerGluIleIleGlnHisGln 56
 123 CTGGAAGCCCTGGAAGTACTATGACATGATGATGATGATGATGATG 172
 57 G1TYrProCysGluGluIlyrGluValAlaThrGluAspGluTyrIle 73
 173 GGATACCCAAATGAGAAATATGAAATGTGATGAAAGATGTTATTTCT 222
 73 uSerValAsnArgIleProArgGlyLeuValGlnProIlySerThrGlyS 90
 223 TGAAGTCAATAGAAATCTTATGGGAAGAAATTCAGGAATTCAGAGCC 272
 90 eArgProValValLeuLeuGlnHisGlyLeuValGlyIlyAlaSerAsn 106
 273 AGAGACCTGTGTGTTTTCACATGCTTGTTCATCAGCCCAAC 322
 107 TrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAsp 123
 323 TGAATTTCCAACTGCGGAGAACAGCTTGCCTTCATCTCGGAGATGC 372
 123 acGlyPheAspValTrpMetClyAsnSerArgGlyAsnAlaTrpSerArg 140
 373 TGGTATGATGATGCTGGTGGCAACACAGACAGAAACACCTGGCCAGAA 422
 140 yHisIlySerThrLeuSerIleAspGlnAspGluPheTrpAlaPheSerTyr 156
 423 GAAACTGTACTATTCACACAGATTCAGTGAATTCCTGGCTTCAGCTTT 472
 157 AspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleGlu 173

473 GATGAATGGCTAAATATGACCTTCAGCCACCAATGACTTCATTTGATA 522
 173 nlySThrGlyGlnGluIlyrValGlyTyrSerGlnGlyThr 190
 523 GAAAGTGGACAGAGACGACTATGTGGGATTCACAGGACCA 572
 190 hMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnIlyS 206
 573 CCATGGTTTATTCCTTTCCACCAATCCACACCTGGGCAAAAGATC 622
 207 LysMetTrpPheAlaLeuAlaProIleAlaThrValIlyHisAlaLys 223
 623 AAACCTTCATGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
 223 rProGlyThrIlyPheLeuLeuLeuProAspMetMetIleGlySlyLeu 240
 673 CCTATTAACAACTAGATTGTTCTCAATCCCTTCCTCAAGTTTAT 722
 240 heGlyLysGluPheLeuTrpGlnThrArgPheLeuArgGlnLeuVal 256
 723 TTGGTGAACAAATATTTCTACCCACACACTTTTGTATCAATTTCTGCT 772
 257 IleTyrLeuGlyGlyGlnValIleLeuAspGlnIleCysSerAsnIle 273
 773 ACTGAAGTGTGCTCCGCTGAGATGCTGATCTCTTGCAGCATGCTCCT 822
 273 tLeuLeuGlyGlyGlyPheAsnThrAsnMetAsnMetSerArgAla 290
 823 ATTATTAATTTGTGATTTGACACTGAAGACTTTAAACGAGTCTGCTGG 872
 290 eValTyrAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeu 306
 873 ATGTGTATCTATCATATTCACAGAGAACTTGTGTTCAAAACATGTTTC 922
 307 HisTrpSerGlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpG 323
 923 CATGGACCCACAGCGCTGTTAGTCTGGGAATTCACACTTATGCTGGGG 972
 323 ySerGluThrIlyAsnLeuGluIlyCysAsnGlnProThrProValArg 340
 973 AAGCCGATTCAGAAATGATGATGATGATGATGATGATGATGATGATG 1022
 340 yArgValAlaArgAspMetThrValProThrAlaMetTrpThrGlyGln 356
 1023 ACAATGTGACAGCCATGATGATGATGATGATGATGATGATGATGATG 1072
 357 AspTrpLeuSerAsnProGluAspValIlyMetLeuLeuSerGluVal 373
 1073 GACCTGTTGGTGCACCCCAAGATGTTGGCTTTGCTTCCAAACATCCC 1122
 373 rAsnLeuIleTyrHisIlyAsnIleProGluTrpAlaHisValAspPhe 390
 1123 CAATCTATTATTCACAAAGAGATTCCTTTTACATCACTTGAACCTTTA 1172
 390 lEtrpGlyLeuAspAlaProHisArgMetTyrAsnGluIleIleHisLeu 406
 1173 TCTGGGCAATGGAGGCCCTCAAGAAATTAACAATGACATTTGTTGATG 1222
 407 MetGlnGlnGlu 411
 1223 ATATCAGAGATAA 1237
 seq_name: /SID8/gcgcdata/geneseq/geneseqn/NA2001.DAT.AAH57420
 seq_documentation_block:
 ID AAH57420 standard; cDNA; 1695 BP.
 XX AAH57420;
 AC
 XX
 DT 10-SEP-2001 (first entry)
 XX
 DE Human stomach cell specific cDNA sequence SEQ ID NO:260.
 XX


```

|||||
1116 CAATCTATTACCAAGAGATCTTTTACAACTGACTGACCTTAA 1165
390 letTpglyLeuaspAlaProHisArgMetIyrAsnGluIleIleHisLeu 406
1166 TCTGGCAAGAGATGCCCCCTCAAGAGTTTACATGACATTTCTTATG 1215
407 MetGlnGlnGlu 411
1216 ATATCAAGAGATAAA 1230
seq_name: /SID58/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF28689

```

```

seq_documentation_block:
ID AAF28689 standard; cDNA; 1308 BP.

```

```

XX AAF28689;
XX
XX 05-APR-2001 (first entry)
XX
XX Human protein HP03372 coding sequence #2.
XX
XX Human; hydrophobic domain; immune deficiency; autoimmune disorder;
XX allergy; tissue growth; regeneration; wound healing; burn; tumour;
XX periodontal disease; thrombolytic condition; haemostatic condition;
XX infection; ss.
XX
XX Homo sapiens.
XX
XX WO200102563-A2.
XX
XX 11-JAN-2001.
XX
XX 16-JUN-2000; 2000WO-IP03943.
XX
XX 02-JUL-1999; 99JP-0188835.
XX
XX (SAGA ) SAGAMI CHEM RES CENT.
XX (PROT-) PROTEGENE INC.
XX
XX Kato S, Kimura T;
XX
XX WPI: 2001-071581/08.
XX P-PSDB: AAB61608.
XX
XX New human proteins with hydrophobic domains, useful for the treatment
XX of immune disorders, tumors, allergic conditions, thrombosis and
XX microbial infection -
XX
XX Claim 4: Pages 122-124; 153pp; English.
XX
XX The present invention relates to human proteins (AAB61608-AAB61617) and
XX their coding sequences (AAF28679-AAF28698). The proteins of the present
XX invention have hydrophobic domains and can be used for the treatment of
XX various immune deficiencies and disorders, such as severe combined
XX immunodeficiency (SCID), multiple sclerosis, rheumatoid arthritis,
XX autoimmune pulmonary inflammation, graft-versus-host disease and
XX Guillain-Barre syndrome. The proteins may also be useful in the treatment
XX of allergic reactions and conditions, such as asthma and in regulation of
XX haematopoiesis or lymphoid cell deficiencies. The proteins may also have
XX utility in compositions used for bone, cartilage, tendon and/or nerve
XX tissue growth or regeneration as well as wound healing and in the
XX treatment of burns. The proteins may be used in the treatment of
XX periodontal disease and in other tooth repair processes. Other uses
XX include treatment of thrombolytic and haemostatic conditions, treatment
XX or prevention of tumours and inhibiting infection by bacteria, viruses,
XX fungi and other parasites.
XX
XX Sequence 1308 BP; 402 A; 246 C; 264 G; 396 T; 0 other;

```

```

alignment_scores:
Quality: 695.00

```

```

Length: 233

```

```

Ratio: 3.717 Gaps: 2
Percent Similarity: 80.258 Percent Identity: 56.223
alignment_block:
US-09-578-063-47 x AAF28689 ..
Align seq 1/1 to: AAF28689 from: 1 to: 1308

```

```

14 ArgMetGluMetTrpLeuLeuIleLeuValAla..... 24
117 TGATCTATGTATGCTTATGACAGAAAGGAACAT..... 153
38 hrLysAlaValaspProGluAlaPheMetAsnIleSerGluIleGln 54
154 .....GCNAACCTGAAGCTTAATATGAATATTAGCCAGATTAATCT 195
55 HisGlnIlyTrpProCysGluGluIuValAlaIleThrGluAspGly 71
196 TACTGGGGTTATCCTTATGAAGACTATGATGACCAAAAGATGGT 245
71 rIleLeuSerValAsnArgIleProArgGlyLeuValGluProIlyst 88
246 TATCCTTGGAATTTATAGGATCCACATGGAAGAGATGCCAGGAGA 295
88 hrGlySerArgProValValLeuLeuGlnIleGlyLeuValGlyAla 104
296 CAGCTCCAAAGCCCTGCTGTATGTCAGCATGCTTAATGCAATGCC 345
105 SerAsnTrpIleSerAsnLeuProAsnAsnSerIleuGlyPheIleuAl 121
346 AGTACTGGATTTGCAACCTGCCCAACAACAGTTTGGCTTCTTCGGC 395
121 aaSPAlaGlyPheaspValTrpMetGlyAsnSerArgGlyAsnAlaTrpS 138
396 AGATAGTGTATGACGCTGTGGGACAGCCGAGAACACTTGCT 445
138 erArgIlySHIlystThrLeuSerIleaspGlnaspGluPheTrpAlaPhe 154
446 CCAGAAAACACCTTAATTTGTCACCGAAATCACCAGAAATACTGGCCTTC 495
155 SerTrpaspGluMetAlaArgPheaspLeuProAlaValIleaspPheI 171
496 AGTTTGATGAGATGCTTAATATGACCTTCCACGCAATCAATTTAT 545
171 eleuGlnIlyThrGlyGlnGluIlystIleTrpValGlyIyrSerGln 188
546 CATAGAGAAACTGACAGAGACGCTTACTAGTGGGCGACACACAG 595
188 IYrTrpMetGlyPheIleAlaPheSerTrpMetProGluLeuAlaGln 204
596 GCACCACCATAGCTTTATAGCATTTCTACAAACCCAGAACTGGCTAAA 645
205 LysIleIlyMetTrpPheAlaLeuAlaProIleAlaThrValIlySHI 221
646 AAGATTAAAGATTTTTCACCTGCTCCAGCTGTACAGTTAAATACAC 695
221 alySerProGlyThrIlyPheLeuLeuLeuProAspMetMetIleIys 237
696 CCAAGTCCTTAAGAAAACCAACACCTTCCAGCGAGTAACTAAG 744
seq_name: /SID58/gcgdata/geneseq/geneseqn/NA2001.DAT:AAF28679
seq_documentation_block:
ID AAF28679 standard; cDNA; 699 BP.
XX AAF28679;
XX
XX 05-APR-2001 (first entry)
XX

```



```

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-227-108-2
seq_documentation_block:
: Sequence 2, Application US/08227108
: Patent No. 5807726
:
: GENERAL INFORMATION:
: APPLICANT: Blanchard, Claire
: APPLICANT: Benicourt, Claude
: APPLICANT: Julien, Jean-Louis
: TITLE OF INVENTION: Recombinant Dog Gastric Lipase
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/227,108
: FILING DATE: 03-APR-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Fannucci, Allan A.
: REGISTRATION NUMBER: 30,256
: REFERENCE/DOCKET NUMBER: 7620-033
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212 790-9090
: TELEFAX: 212 869-8864/9741
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1137 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1137
:
: US-08-227-108-2
:
: Alignment_scores:
: Quality: 1141.50 Length: 372
: Ratio: 3.624 Gaps: 3
: Percent Similarity: 84.677 Percent Identity: 56.452
:
: Alignment_block:
: US-09-578-063-47 x US-08-227-108-2
:
: Align seg 1/1 to: US-08-227-108-2 from: 1 to: 1137
:
37 ProThnLysAlaValAspProGluAlaPheMetAsnIleSerGluIleI 53
||||| :::|||||::: |||||||:::|||||
19 CCCACA.....AACCTGAAGTGCACATGAATATAGTCAGATGAT 59
53 eGlnHsGlnGlyTyrProCysGluGluTyrGluValAlaThrGluAspG 70
| ::: |||||||:::|||||:::|||||:::|||||
60 CACCTACTGCGGATGCCAGCTGAGCAATATGCAAGTTGTGCAGCAGAGAG 109
70 lYtYrILleuSerValAsnArgIleProArgGlyLeuValGlnProLys 86

```

```

1007 TTTCGAAGCTCCCAATTCATTACACAGAGAGATTCCTCTTACAAT 1056
386 HisValAspHeIleTrrpGlyLeuAspAlaProHisArgMetTyraSngl 402
||||:||||:||||:||||:||||:||||:||||:||||:||||:
1057 CACTTGACTTATCTAGTGGCCATGATGATGCCCTCAGCGGTTTCAATGA 1106
402 uIleIleHisLeuMet 407
||||:||||:||||:||||:||||:||||:||||:||||:||||:
1107 AATTGTTTCATGATG 1122

seq_name: /sgn2_6/ptodata/2/lna/5B_COMB.seq:US-09-073-674-2

seq_documentation_block:
; Sequence 2, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Julien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM-PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crisey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-DI-66-TWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1137 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1137
; US-09-073-674-2

alignment_scores:
Quality: 1141.50 Length: 372
Ratio: 3.624 Gaps: 3
Percent Similarity: 84.677 Percent Identity: 56.452

alignment_block:
US-09-578-063-47 x US-09-073-674-2 ..
Align seg 1/1 to: US-09-073-674-2 from: 1 to: 1137

37 ProThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIle 53
||||| :|||||:|||||:|||||:|||||:|||||:|||||:|||||
19 CCACAC.....AACCTGAAGTACCATGATATTAAGTCAGATGAT 59
53 eginInhSgInGlyTyrProCysGluGluTyrGluValAlaThrGluAspG 70

```

1
60 CACCTACTGGGGATPACCCAGCTGAGGAATATGAAATTGTGACCGGAAGAC 109
70 LTYTTLLEuSerValAsnArgIleProArgIlyLeuValGlnProLys 86
110 GTTATATCCTTGGGATGACAGAAATTCCTTATGGAGGAAATAATTCAGAG 159
87 LysThrGlySerArgProValValLeuLeuGlnHisGlyLeuValGlyIyl 103
160 AATATAGCGCCGAGACCTGTTGCATTTTGTGCACACGGTTTCTCGCAATC 209
103 yAlaSerAsnIlePLeSerAsnLeuProAsnAsnSerLeuGlyPheIleL 120
210 AGCCACAACATCGATCTCCAACTCCGCCAACACAGCCCTGGCCCTTATCC 259
120 euAlaAspAlaGlyPheAspValTrrPmetGlyAsnSerArgIlyAsnAla 136
260 TGGCGGACGCCGGGATGACACGCTGTGGCTGGGACACGAGGGGACACAC 309
137 TrrSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrrAl 153
310 TGGCGCAGAGGAATCTGTACTACGCCGCCACTCCGTGGAATTCGGGCG 359
153 aphSerTrrAspGluMetAlaArgPheAspLeuProAlaValIleAsnP 170
360 TTTGAGCTTTACGAGATGCGTAATATGACCTTCCGCCACCATTTGACT 409
170 heileLeuGlnLysThrGlyGlnGlnLysIleTyrTrrValGlyTrrSer 186
410 TCATCTTGAACAAAAAGGAGACGGACAGCAAGCTACACTACTGTGGCCATTCC 459
187 GlnGlyTrrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
460 CAGGCAACACCATTTGTTTCAATTCGCTTTCACCAATCCAACTGAGCTGG 509
203 agLysIleLeuLysMetTrrPheAlaLeuAlaProIleAlaThrValLysH 220
510 GAAACGGATCAAAACCTTCTATGCAATTCGCTCCCGTTGCCACCGTAAAGT 559
220 lSAlaLysSerProGlyThrLysPheLeuLeuLeuProAsnMetIleL 236
560 ACACGGAACCCCTGTTAAACAAACATCATGCTGCTCCCTTCGTCCCTTC 609
237 LysGlyLeuPheGlyLysLysGluPheLeuTrrGlnThrArgPheLeuAr 253
610 AAGCTTATATTTGGAACAACAATATTC..TTCACACACACACTTCTTGA 656
253 gGln..LeuValIleTrrLysCysGlyGlnValIleLeuAspGlnIleC 269
657 TCAAATTTCTCCGCCACCGAGGTATGCTCCCGCAGAGCGGTGATCTCTCT 706
269 ysserAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnMetAsn 285
707 GCAGCAACGCCCTGTTTATCATTTGTGGATTTGACACTTATGAACCTTGAA 756
286 MetSerArgAlaSerValTrrAlaIleHisThrLeuAlaGlyThrSerVa 302
757 ATGATTCCTGTTGAGTGTGATCTGCACATAATCCAGAGGAACATCGGT 806
302 lGlnAsnIleLeuHisTrrPserGlnAlaValAsnSerGlyGluLeuArgA 319
807 TCAGAACGTGCTCCACTGTGCTCCAGAGCTTAAAGTCTGGGAACTTCCAG 856
319 lApheAspTrrGlySerGluTrrThrLysAsnLeuGlnLysCysAsnGlnPro 335
857 CTTTGTACTGGGGAAAGCCAGTTCAGAAATGATGTATGCATATCTACGAGC 906
336 ThrProValArgTrrArgValAlaArgAspMetThrValProThrAlaMet 352
907 ATGCCTCCCTACTACAACCTGCACAGACATGACATGTGTCCAATGCACAGTG 956
352 PThrGlyGlyGlnAspTrrPLeuSerAsnProGluAspValLysMetLeu 369

```

957 GAACGCTGGCAAGCATCTTGCTGGCGCCAGCCTCAAGATGTTGACCTTTGCG 1006
369 eUSErclUvAlYtHrAsnLeuIleTyHIsLySAsnIleProGluTrPAla 385
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1007 TTTCCAGAGCTCCCAATCTCATTTTACCACAGGAGATTCCTCTTACAT 1056
386 HIsValAspHeIleTrPglYLeAsPAlaProHIsArGmeTyYrAsnG1 402
1057 CATTGGACTTATATCTGGCCATGATGATGCCCTCAAGCGGTTTCAATGA 1106
402 uIleIleHIsLeuMet 407
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
1107 AATTGTTTCATGATG 1122

seq_name: /cgn2_6/plodata/2/lna/5A_COMB.seq:US-08-227-108-4

seq_documentation_block:
: Sequence 4, Application US/08227108
: Patent No. 5807726
: GENERAL INFORMATION:
: APPLICANT: Blanchard, Claire
: APPLICANT: Benicourt, Claude
: APPLICANT: Junien, Jean-Louis
: TITLE OF INVENTION: Recombinant Dog Gastric Lipase
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/227,108
: FILING DATE: 03-APR-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Fanucci, Allan A.
: REGISTRATION NUMBER: 30,256
: REFERENCE/DOCKET NUMBER: 7620-033
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212 790-9090
: TELEFAX: 212 869-8864/9741
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1140 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1140
: US-08-227-108-4

alignment_scores:
Quality: 1141.50 Length: 372
Ratio: 3.624 Gaps: 3
Percent Similarity: 84.677 Percent Identity: 56.452

alignment_block:
US-09-578-063-47 x US-08-227-108-4 ..
Align seg 1/1 to: US-08-227-108-4 from: 1 to: 1140

37 ProHtLySAlaValAspProGluAlaPheMetAsnIleSerGluIle1 53

```



```

|||||
22 CCCACA.....AACCTGAGTACCATGATATTAAGTCAGATGAT 62
53 eglhlsigllytrpocysglutrygluvalalaThrcluasp 70
63 CACCTACGTGGGATACCACCTGAGNAATGAAGTTGTGACCGAAGAC 112
70 lYtYrIleuSerValAsnArgIleProArgGlyLeuValGlnProlys 86
113 GTTATATCCTGGGATCGACAGAAATCTTATGGGAGAAAATTCAGAG 162
87 LysThrGlySerArgProValIleuLeuGlnHisGlyLeuValGly 103
163 AATATAGGCGGAGACCTGTGATTTTCCACACGCTTGCCTGCATC 212
103 yAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIle 120
213 AGCCACAAACTGGATCTCCAACTGCCCAACAGAGCTGCGCTTATCC 262
120 euAlaAspAlaGlyPheAspValTyrMetClyAsnSerArgGlyAsnAla 136
263 TGCCCGACGCGGGGTACGACGTGTGGGGAACAGACAGCGGCAACACC 312
137 TrpSerArgLysHisLysThrIleuSerIleAspGlnAspGluPheTrpAl 153
313 TGGGCGACGAGGAATCTGACTACTGCGCCGACTCCGTGCAATTTGTGGC 362
153 aPheSerTrpAspGluMetAlaArgPheAspLeuProAlaValIleAsn 170
363 TTTACACTTGAACGAGATGCGCTAATATGACTTCCCGCCACCATTGACT 412
170 helIleuGlnLysThrGlyGlnGlyLysIleTyrValGlyTyrSer 186
413 TCATCTTGAAGAAACGGACAGACAGACATACACTGTCGCATTC 462
187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeu 203
463 CAGGGACACCATGTTTTCATCGCCTTTCCACCAATCCCAAGCTGGC 512
203 aGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThrValLys 220
513 GAACCGATCAAAACCTCTATCATGATCCGCTTCCACCGTGAAGT 562
220 lAlaLysSerProGlyThrLysPheLeuLeuProAspMetMetIle 236
563 AACCGAAACCTGTTAAACAACTATGCTGCTCCCTTCCTTCCTTC 612
237 LysGlyLeuPheGlyLysLysGluPheLeuTyrGlnThrArgPheLeu 253
613 AAGCTATATTTGGAACAAATATTC...TACCCACACACTTCTTTGA 659
253 geln..LeuValIleTyrLeuGlyGlnValIleLeuAspGlnIleC 269
660 TCATATTTCTCGCCACCGAGATGCTCCGCGAGAGCTGATCTCTCT 709
269 ySerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnMetAsn 285
710 GCAGCAACGCCCTGTTTATCATTTGTGATTTGACACTATGAACCTTGA 759
286 MetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSer 302
760 ATGAGTCGCTGGATGATGATGTCACACTATATCCAGCGAGACATTCG 809
302 lGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyLeuLeuArg 319
810 TCAGAACGTCCTCACTGCTCCAGGCTGTAAAGTCTGGGAAGTCCAAAG 859
319 lAPheAspTrpGlySerGluThrLysAsnLeuGlnLysCysAsnGlnPro 335
860 CTTTGAAGTGGGAGCCAGTTCAAGACATGATGACATATCATCAGAGC 909
336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMet 352

```

```

910 ATGCTTCCTACTACAACTGACAGACATGATGTCCTCAATCCGAGTGTG 959
352 pThrGlyGlyLysTrpLeuSerAsnProGluAspValLysMetLeu 369
960 GAACGGTGGACAGACTGCTGGCGACCCCTCAACAGATGTGACTTTGC 1009
369 euSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAla 385
1010 TTTCAAGCTCCCAATCTCATTTACACAGAAATTCCTCTTCAAT 1059
386 HisValAspPheIleTrpGlyLeuAspAlaProHisArgMetLysAsn 402
1060 CACTTGGACTTATCTGGGCAATGATGCCCTCAAGCGGTTTACATGA 1109
402 uIleIleHisLeuMet 407
1110 AATTGTTTCATGATG 1125
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-09-073-674-4

```

seq_documentation_block:

```

: Sequence 4, Application US/09073674
: Patent No. 5998189

```

GENERAL INFORMATION:

```

: APPLICANT: Blanchard, Claire
: APPLICANT: Benicourt, Claude
: APPLICANT: Junien, Jean-Louis
: TITLE OF INVENTION: Recombinant Dog Gastric Lipase
: NUMBER OF SEQUENCES: 21
: CORRESPONDENCE ADDRESS:
: ADDRESS: Warner-Lambert Company
: STREET: 2800 Plymouth Road
: CITY: Ann Arbor
: STATE: Michigan
: COUNTRY: U.S.A.
: ZIP: 48105

```

COMPUTER READABLE FORM:

```

: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/073.674
: FILING DATE:

```

CLASSIFICATION:

```

: ATTORNEY/AGENT INFORMATION:
: NAME: Clissey, Todd M.
: REGISTRATION NUMBER: 37,807
: REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 734 622-7530
: TELEFAX: 734 622-1553
: TELEX:

```

INFORMATION FOR SEQ ID NO: 4:

```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 1140 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1140
: US-09-073-674-4

```

alignment_scores:

```

: Quality: 1141.50 Length: 372
: Ratio: 3.624 Gaps: 3
: Percent Similarity: 84.677 Percent Identity: 56.452

```

alignment_block:

```

US-09-578-063-47 x US-09-073-674-4

```

```

860 CTTTGACTGGGGAAGCCAGTTCAGAAACGATGCATCATCAGAGC 909
336 ThProValArgYrArGValAraGasPmethrValProHralametr 352
    ||| ||||| ||||| ||||| |||||
910 ATGCCCTCCCTACCAACAACCTGCACGAGATGCATGTGCCAATCGAGTGTG 959
352 pThrgIyGlAspPrpleuseArsanProGluAspVallysmetLeuL 369
    |:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
960 GAACGGTGGCAAGCACTGCTGTGCCGACCCTCACGATGTACCTTTTGC 1009
366 euserGluValThrAsnLeuIleTyRHisLysAsnIleProGluTrpAla 385
    |||||::: ||||| ||||| ||||| ||||| :::
1010 TTTCCAAGCTCCCCAATCTCATTTACACAGAAGATTCTCTTACAT 1055
386 HisValAspHeiletrpGlyLeuAspAlaProHisArgmetYrAsnGI 402
    |||||:::|||||:::|||||:::|||||:::|||||:::|||||
1060 CACTTGACCTTATCTATCTGGCCATGATGATGCCCTCAACGGCTTCAAAGA 1109
402 uillelleHisLeumet 407
    |||||:::|||||
1110 AATTGTTCCATGATG 1125

seq_name: /cgn2_6/plodata/2/ina/5A_COMB.seq:US-08-227-108-6

seq_documentation_block:
; Sequence 6, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1146 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
US-08-227-108-6

alignment_scores:
Quality: 1141.50 Length: 372
Ratio: 3.624 Gaps: 3
Percent Similarity: 84.677 Percent Identity: 56.452

alignment_block:
```

US-09-578-063-47 x US-08-227-108-6
 Align seq 1/1 to: US-08-227-108-6 from: 1 to: 1146

37 ProthlyslalValasProgluaAlaPheMetAsnIleSerGluIle1 53
 19 CCCACA.....AACCTGAAGTGCATGATATTAAGTCAGATGAT 59
 53 eglInHISglnglyTyTProCysgluGluValAlaThGluaspG 70
 60 CACCTGAGGGATACCACCTGAGGAATGAGTGTGACCGAAGACG 109
 70 LyTylIleuSerValAsnArgIleProArgGlyLeuValGlnProLys 86
 110 GTTATATCCTTGGGATCGACAGAAATTCCTTATGGAGAAAAATTCAGAG 139
 87 LysThylGlySerArgProValValLeuLeuGlnHISglYleuValGly 103
 160 AATATAGCGCGGAGACCTGTTGCAATTTTGCACACAGGTTGCTGCAGATC 209
 103 yAlaSerAsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleL 120
 210 AGCCACAACCTGATCTCCAACTGCCCAACACAGCCTGCTGCCTCATCC 259
 120 euAlaSpAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136
 260 TGCCCGACGCGGGTACGAGCTGTGCGGGGAACAGCAGGGGCAACACC 309
 137 TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGlnPheTrpAl 153
 310 TGGGCGAGAGAGAACTGTACTACTGCGCCGACCTCCGTGCAATTCGTGGC 359
 153 aPheSerTrpAspGluMetAlaArgPheAspLeuProAlaValIleAsn 170
 360 TTTCACCTTGAAGAGATGGCTAAATATGACCTCCCGCACCATGACAT 409
 170 heLeuGlnLysThrGlyGlnGluLysIleTyTyrValGlyTyrSer 186
 410 TCATCTGGAAGAAACGAGACGAGACAGTACACTACCTGGCCATTC 459
 187 gInglyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAl 203
 460 CAGGGACACCATGTTTCATCGGCTTTCCACCAATCCCAAGCTGGC 509
 203 aGlnLysIleLysMetTrpPheAlaLeuAlaProIleAlaThrValLysH 220
 510 GAACGAGCAAAACCTCTATGACATTCCTCCGTCACCGCTGAAG 559
 220 lAlaLysSerProGlyThrLysPheLeuLeuLeuProAspMetMetIle 236
 560 ACACCGAAACCTGTTAAACAACTCATCTCTCCCTTCCTTCCTTC 609
 237 LysGlyLeuPheGlyLysLysGluPheLeuTyTyrGlnThrArgPheLeu 253
 610 AAGCTTATATTGGAACAAATATTC...TACCCACACCACTTCTTTGA 656
 253 gGln..LeuValIleTyTyrLeuCysGlyGlnValIleLeuAspGlnIleC 269
 657 TCAATTTCTCGCCACCGAGTATGCTCCCGGAGAGCGTGAATCTCTCT 706
 269 ySerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnMetAsn 285
 707 GAGACGACCCCTGTTATCATTTGTTGACTTGAACATVGAACCTTGAAC 756
 286 MetSerArgAlaSerValTyTyrAlaAlaHisThrLeuAlaGlyThrSer 302
 757 ATGAGTCGCTTGATGTATCTGCACATATATCCAGCAGAAACATCGGT 806
 302 lGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyLuleuArg 319
 807 TCAGAACGTGCTCACCTGCTCCAGGCTGTTAAGTCTGGGAGATTCCAAG 856
 319 lAPhaSPTrpGlySerGluThrLysAsnLeuGlnLysCysAsnGlnPro 335

857 CTTTGTGACTGGGGAAGCCAGTTCCAGAACATGATGACATCATCAGACG 906
 336 ThrProValArgTyTyrArgValArgAspMetThrValProThrAlaMet 352
 907 ATGCGTCCCTACACTACAACTGACAGACATGATGTCGCAATCGCAGTGTG 956
 352 pThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeu 369
 957 GAACGGTGGCAACGACTGCTGGCCGACCTCCACGATGTGACTTTCG 1006
 369 euSerGluValThrAsnLeuIleTyTyrHisLysAsnIleProGluTrpAla 385
 1007 TTTCAGAGCTCCCAATCTCATTTACCCACAGAAATTCCTCTTCAAT 1056
 386 HisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyTyrAsn 402
 1057 CACTTGGACTTATCTGGGCGCATGATGCCCTCACAACGGGTTTACATGA 1106
 402 uIleIleHisLeuMet 407
 1107 AATGTGTTCCATGATG 1122

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-09-073-674-6

seq_documentation_block:

Sequence 6, Application US/09073674
 Patent No. 5998189

GENERAL INFORMATION:

APPLICANT: Blanchard, Claire
 APPLICANT: Benicourt, Claude
 APPLICANT: Junien, Jean-Louis
 TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Warner-Lambert Company
 STREET: 2800 Plymouth Road
 CITY: Ann Arbor
 STATE: Michigan
 COUNTRY: U.S.A.
 ZIP: 48105

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/073,674
 FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
 NAME: Crissey, Todd M.
 REGISTRATION NUMBER: 37,807
 REFERENCE/DOCKET NUMBER: 5072-DI-66-TMC
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 734 622-7530
 TELEFAX: 734 622-1553

TELEX:

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1146 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 US-09-073-674-6

alignment_scores:

Quality: 1141.50 Length: 372
 Ratio: 3.624 Gaps: 3
 Percent Similarity: 84.677 Percent Identity: 56.452

alignment_block:
US-09-578-063-47 x US-08-227-108-1 ..

Align seg 1/1 to: US-08-227-108-1 from: 1 to: 1528

```

37 ProthrlyslalavalaspProgluAlaphemetaInleserGluIleI 53
   |||||
   |||||
19 CCCACA.....AACCTGAAGTACCATGATATATAGTCACATCAT 59
   |||||
   |||||
53 eGlnHisGlnGlyTyrProCysGluGluTyrGluValAlaThrGluAsp 70
   |||||
   |||||
60 CACCTACGCGGATACCCAGCTGAGGAATATGAACTTGACCGAAGAGC 109
   |||||
   |||||
70 IYTYRILEuSerValAsnArgIleProArgGlyLeuValGlnProlys 86
   |||||
   |||||
110 GTTATATCTCTGGGATCGACAGAAATCTCTATGAGAGAAAAATTCAG 159
   |||||
   |||||
87 LysThrGlySerArgProValValLeuLeuGlnHisGlyLeuValGly 103
   |||||
   |||||
160 AATATAGCGCGAGACCTGTTGCATTTTGCACACGGTTGCTGCATC 209
   |||||
   |||||
103 YAlAsenTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIle 120
   |||||
   |||||
210 AGCCCAAACTGGATCTCCAACTGCCCAACACAGCTGGCTTCATCC 259
   |||||
   |||||
120 euAlaAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAla 136
   |||||
   |||||
260 TGGCCGACCGCGGTACAGCGTGTGGGAGACAGAGGGGCAACACC 309
   |||||
   |||||
137 TrpSerArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrp 153
   |||||
   |||||
310 TGGCCGACGAGGAATCTGACTACCGCCGACCTCGTGCATTTCTGG 359
   |||||
   |||||
153 aPheSerTyrAspGluMetAlaArgPheAspLeuProAlaValIleAs 170
   |||||
   |||||
360 TTTGAGCTTTGACGAGATGGCTAAATATGACTCTCCCGCACCATG 409
   |||||
   |||||
170 heIleLeuGlnLysThrGlnGlnLysIleTyrTyrValGlyTyrSer 186
   |||||
   |||||
410 TCATTTTGAAGAAACGGGACAGCAAGCTACACTACCTTGCCCATTC 459
   |||||
   |||||
187 GlnGlyThrThrMetGlyPheIleAlaPheSerThrMetProGluLeu 203
   |||||
   |||||
460 CAGGACACCATGTTTTCATCGCTTTCCACCAATCCCAACCTGGC 509
   |||||
   |||||
203 aGlnLysIleLysMetTyrPheAlaLeuAlaProIleAlaThrValLys 220
   |||||
   |||||
510 GAAAGGATCAAAACCTTCTATGCTAGCTCCCTGCGACACCTGAAG 559
   |||||
   |||||
220 IAlaLysSerProGlyThrLysPheLeuLeuProAspMetIle 236
   |||||
   |||||
560 ACACCGAAACCTGTTAAACAACTCATGCTGCTCCTCGTCTCTTC 609
   |||||
   |||||
237 LysGlyLeuPheGlyLysGluPheLeuTyrGlnThrArgPheLeuAr 253
   |||||
   |||||
610 AAGCTATATTTGGAACAAATATTC..TACCCACACACTTCTTTGA 656
   |||||
   |||||
253 gGln...LeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIle 269
   |||||
   |||||
657 TCAATTTCTCGCACGAGGATGCTCCCGCAGAGAGGTGATCTCTCT 706
   |||||
   |||||
269 YSerAsnIleMetLeuLeuGlnGlyPheAsnThrAsnMetAsn 285
   |||||
   |||||
707 GCAGCAAGCCCTGTTATCATTTGTGATTTGACACTATGACCTTGAC 756
   |||||
   |||||
286 MetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSer 302
   |||||
   |||||
757 ATGAGTCTGCTGGATGTGATCTGCACATATCCAGCAGGAACATCGT 806
   |||||
   |||||
302 IGlAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyIleLeuAla 319
   |||||
   |||||
807 TCAGAACGTCTCACCTGCTCCAGGCTTTAAGTGTGGGAATTCACAG 856

```

```

319 IAlaPheAspTrpGlySerGluThrLysAsnLeuGlnLysCysAsnGlnPro 335
   |||||
   |||||
857 CTTTGTACTGGGGAGCCAGTTCAGACATGATGACATCATCATCAGAGC 906
   |||||
   |||||
336 ThrProValArgTyrArgValArgAspMetThrValProThrAlaMetTr 352
   |||||
   |||||
907 ATGCTCTCCCTACTACACCTGACAGACATGATGCGCAATCGCAGTGTG 956
   |||||
   |||||
352 PThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeu 369
   |||||
   |||||
957 GAAAGGTGACAGACCTTGTGCGCCGACCTGCACATGTTGACCTTTGC 1006
   |||||
   |||||
369 euSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAla 385
   |||||
   |||||
1007 TTTCCAACTGCCCAATCTCATTTACCAAGAGAAATCTCTTACAT 1056
   |||||
   |||||
386 HisValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnG 402
   |||||
   |||||
1057 CACTTGACCTTATCTGGGCGCATGATGATGCCCTCAAGCGTTTACAAAGA 1106
   |||||
   |||||
402 uIleIleHisLeuMet 407
   |||||
   |||||
1107 AATGTTTCCATGATG 1122

seq_name: /cgn2_6/ptodata/2/1na/5b_COMB.seq:US-09-073-674-1

seq_documentation_block:
; Sequence 1, Application US/09073674
; Patent No. 5998189
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junten, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Warner-Lambert Company
; STREET: 2800 Plymouth Road
; CITY: Ann Arbor
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 48105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0., Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,674
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Crissey, Todd M.
; REGISTRATION NUMBER: 37,807
; REFERENCE/DOCKET NUMBER: 5072-D1-66-TMC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 734 622-7530
; TELEFAX: 734 622-1553
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1528 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
US-09-073-674-1

alignment_scores:
Quality: 1141.50 Length: 372
Ratio: 3.624 Gaps: 3

```



```

14869 CCATTGTGATGCTACATGCGCTTTGCAAGACGATGCTGCATCC 14918
110 AsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAspAlaGlyPheAs 126
14919 AAT...GGTAGAAATCTCTGGCATATTTCTGATCAATCCGGTAGCA 14965
126 PValTrpMetGlyAsnSerArgGlyAsn.....AlaTrpSerArg. 139
14966 CATATGCTTAGGAATTAACAGATGGGGTTTAGCCCGAATGGAACGAAG 15015
140 ..LysHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSer 155
15016 CGAAAGTACCCACTACTGCT.....TCCAGGTGGAGCTGGAC 15053
156 TyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLe 172
15054 CTTCGGAGATGGTTAAGTACGATCTGACCTTTGGATTGATACCGGTGT 15103
172 uGlnLysThrGlyGlnGlyLysIleTyrValGlyTyrSerGlnGlyT 189
15104 AGCTAAGACGACGATTGAAAAGCTTACTTGTGATCTCGCATCTCAGGCA 15153
189 hTrpMetGlyPheIleAlaPhe..... 196
15154 CTACACAGGGGTTATGGGCTTGTCACAGCAAGATATTTTCCCTGCC 15203
197 .....SerThrMetProGluLeuAlaGlnLysIleLysMetTyrPhe 210
15204 GGTTCGGATCTAAGAAATCTTTTTCATCTTAAGATCGCAAACTATAT 15253
210 eAlaLeuAlaProIleAlaThrValHisAlaLysSerProGlyThrL 227
15254 TGCCCTGGCCCC.....GCAGTGTATCCTGGTCCC. 15284
227 ySPheLeuLeuProAspMetIleLys..... 237
15285 ....TTACTTAACGAAATTTGTTGTTAAGCTTATGACAAGGAATC 15329
238 .....GlyLeuPheGlyLysLysGluPheLeuTyrGlnThrArgPhe 251
15330 GAAAATCCCTGGTCTTGTGTAACGAGCTTTTGTGAGATAATGATAT 15379
251 eLeuArgGlnLeuValIleTyrLeuCySGlyGlnValIleLeuAspGlnI 268
15380 TGTAAGAACTTGTGCGTT.....GGTAGGCTGTGTTCTCTCTTGG 15420
268 leCySerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnMet 284
15421 TTTGTTACACATCTTCAATTACCGTTGATGGAACGATACCTTTGG 15470
285 AsnMetSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThySe 301
15471 GATACCGCATTAAGAGATGCCATTCTCTGTTTCGCCAGTCCATGTTTC 15520
301 rValGlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuA 318
15521 AGTGAG...TTGATGCAATGCTGCTGTACCCGACCCCAACGAAGGTA. 15566
318 rGAlaPheAspTrpLeuSerGluThrLysAsnLeuGluLysCysAsnGln 334
15567 ..AGTTTAAATTTGGTCCCATTAAGATG.....TTC 15596
335 ProThrProValArgTyrArgValArgAspMetThrValProThrAla. 350
15597 CCCGCAATGTAAAGTGTGTTTACAGACCATCAAGGCCCAAAATATCTA 15646
351 .MetTrpThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysM 367
15647 CTGTGTTTTCCAAGCAAGATAGTTG.....GTGACGAGAGAAA 15687
367 etLeuLeuSerGluValThrAsnLeuIleTyrHisLysAsn..... 380
15688 GACTAATCAATTCCTCAATGCTGAGTGCATGCAATGCACTACAGATC 15737

```

```

381 .....IleProGluTrpAlaHisValAspPheIleTrpGlyLeuAspAl 395
15738 TGATCATTTGATAGTATGCCCATTTGATGCTTATGGGACATGATGT 15787
395 aProHisArgMetTyrAsnGluIleLe 404
15788 CATAGAGCAATTTGTAACCACTTTTA 15815

seq_name: /cgn2.6/p1oddata/2/lna/5A_COMB.seq:US-08-751-782-2
seq_documentation_block:
; Sequence 2, Application US/08751782
; Patent No. 5821352
; GENERAL INFORMATION:
; APPLICANT: Heintz, Nathaniel
; APPLICANT: Gubbay, Jonathan
; APPLICANT: Skinner, Michael
; TITLE OF INVENTION: A cDNA library prepared during
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/751,782
; FILING DATE: 18-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-190
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 178 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; IMMEDIATE SOURCE:
; CLONE: 10.2
; US-08-751-782-2

alignment_scores:
Quality: 155.00 Length: 50
Ratio: 3.780 Gaps: 0
Percent Similarity: 82.000 Percent Identity: 56.000

alignment_block:
US-09-578-063-47 x US-08-751-782-2 ...
Align seg 1/1 to: US-08-751-782-2 from: 1 to: 178

139 ArgLysHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSe 155
|||||
7 AGGAAACATGTGGCGCTTAGACCAAGATTTCTAAGAAATTTGGGATTTTG 56
155 rTyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleL 172

```

```

172 euglnlsthrglynglulyslleetyrttyvalgityrsergngly 188
177 TGAATGAAACAGACAAACAAATTTACTATATGGCCATTCGCCAGGC 156
107 TGAATGAAACAGACAAACAAATTTACTATATGGCCATTCGCCAGGC

```

seq_name: /cgn2_6/plodata/2/1na/5A.COMB.seq:US-08-925-171-2

seq_documentation_block:

```

; Sequence 2, Application US/08925171
; Patent No. 5928871
; GENERAL INFORMATION:
; APPLICANT: Heintz, Nathaniel
; APPLICANT: Gubbay, Jonathan
; APPLICANT: Skinner, Michael
; TITLE OF INVENTION: A CDNA Library Prepared during
; TITLE OF INVENTION: Regression of Rat Prostate and Methods of Use Thereof
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; STREET: Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/925,171
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-190 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 178 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; IMMEDIATE SOURCE:
; CLONE: 10.2
; US-08-925-171-2

```

alignment_scores:

```

Quality: 155.00 Length: 50
Ratio: 3.780 Gaps: 0
Percent Similarity: 82.000 Percent Identity: 56.000

```

alignment_block:

US-09-578-063-47 x US-08-925-171-2 ..

Align seg 1/1 to: US-08-925-171-2 from: 1 to: 178

```

139 ArGlySHisYthrIleuSerIleasGlnAspGluPhetrlAlaphe 155
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7 AGGAACATGTGGCCCTAGACCCAGATTCGAAAGATTTGGGATTTAG 56
155 rTyAspGluMetAlaArgPheAspLeuProAlaValIleAsnPhetrl 172
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
57 TTTTATGAAACAATAGATAGACACCTCCAGCATCATTTATTTTCATTC 106

```

```

172 euglnlsthrglynglulyslleetyrttyvalgityrsergngly 188
177 TGAATGAAACAGACAAACAAATTTACTATATGGCCATTCGCCAGGC 156
107 TGAATGAAACAGACAAACAAATTTACTATATGGCCATTCGCCAGGC

```

seq_name: /cgn2_6/plodata/2/1na/5A.COMB.seq:US-08-484-105-21

seq_documentation_block:

```

; Sequence 21, Application US/08484105
; Patent No. 5589341
; GENERAL INFORMATION:
; APPLICANT: STILLMAN, Bruce
; APPLICANT: BELL, Stephen P
; APPLICANT: KOBAYASHI, Ryuji
; APPLICANT: RINE, Jasper
; APPLICANT: FOSS, Margit
; APPLICANT: McNALLY, Francis J
; APPLICANT: LAURENSEN, Patricia
; APPLICANT: HERSKOWITZ, Ira
; APPLICANT: LI, Joachim J
; APPLICANT: GAVIN, Kimberly
; TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ELEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,105
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman Ph.D., Richard Aron
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1676 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 13..1302
; US-08-484-105-21

```

alignment_scores:

```

Quality: 105.00 Length: 278
Ratio: 0.802 Gaps: 13
Percent Similarity: 47.122 Percent Identity: 21.942

```

alignment_block:

US-09-578-063-47 x US-08-484-105-21 ..

Align seg 1/1 to: US-08-484-105-21 from: 1 to: 1676

```

85 ProLysLysThrglySerArgProValIleuLeuGlnHscglyLeuVa 101
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
85 CAGAAAAAAGAGATCCAGCA.....CAGAAAAAGAC 116

```



```

101 IGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsnAsn..... 114
117 GAATGGAAGAAAGAAATGCTTTCAGAAATTTGCATCAATTTAGAGAAG 166
115 .....SerLeuGlyPhe.....IleuAlaAspAla 123
167 ATTTGGAACAACTGGCTTCAGAGATGAAGTGTATCAATGGCTCAATCA 216
134 GlyPheAspValTrp.....MetGlyAsnSer.....Arg91 134
217 GCAATCGAAATTAATTAATGCAAGAAATGGGCTCAGACGAATGAA 266
134 yAsnAlaTrpSerArgGlySHisLys.....ThrLeuS 145
267 TAATCGGAATCCCTCGTGGAGAGCTGCGAAATGGAATGAACTGAAG 316
145 eTlleAspGlnAspGluPheTrpAlaPheSerTyrAspGluMetAla 161
317 AAATTGAGAGACGATGATGATCAATGATGATGATGATGATGATGATG 366
162 PheAspLeuProAlaValIleAsnPheIleLeuGlnLysThrGlyGln 178
367 TGTGATCTCCCTGAGCTCGAAATATATATACCAAAAAAGATACACGA 416
178 uLysIleTyrTyrValGlyTyrSerGlnGlyThrMetGlyPheIle 195
417 ATTCGAAAACGATGAGATCTCGCGATGATGATGATGATGATGATG 459
195 laPheSerThrMetProGluLeuAlaGlnLysIleLysMetTyrPheAla 211
460 .....AAATGGAAGCTTAACTGATGATGATGATGATGATGATGATG 480
212 leuAlaProIleAlaThrValLysHisAlaLysSerProGlyThrLys 228
481 GCTGGA..... 486
228 eleuLeuLeuProAspMetMetIleLysGlyLeuPheGlyLysGluP 245
487 .....TTTAATATCTCTTTCACGCGTCTCGTTCGGAAGCGTGAG 526
245 heLeuTyrGlnThrArgPheLeuArgGlnLeuValIleTyrLeuGly 261
527 TTCTC.....ACAGATTTGGAATGAGCTATCCGATTTATCATATATG 570
262 GlnValIleLeuAspGlnIleCysSerAsnIleMetLeuLeuGlyG 278
571 AGAGTGAATGCACGGAAGATGGCTCAATGTAAGTTCTCTTGAGAC 620
278 yPheAsnThrAsn.....AsnMetAsnMetSerArgAlaSerValT 292
621 TATCATATGAAATATGAAAGCTGAATGTGATGTGAAAGAGCG..... 663
292 yTAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeuHis 308
664 .....CAATCTACGATTTGTTGG 681
309 SerGlnAlaVal.....AsnSerGlyGluLeu..... 317
682 GCTCATCTATTCGCAAGAAATGAAATGACCAACAGTTGATTTATCAT 731
318 .....ArgAlaPheAspTrpGlySerGlu 325
732 TGATATATATGAAAGCTCTGATGTGGAAGTGAAT 765

```

seq_name: /cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-484-106-21

seq_documentation_block:

; Sequence 21, Application US/08484106

; Patent No. 5614618

; GENERAL INFORMATION:

; APPLICANT: STILLMAN, Bruce

; APPLICANT: BEILL, Stephen P

```

APPLICANT: KOBAYASHI, Ryuji
APPLICANT: RINE, Jasper
APPLICANT: FOSS, Margit
APPLICANT: McNALLY, Francis J
APPLICANT: LAURENSEN, Patricia
APPLICANT: HERSKOWITZ, Ira
APPLICANT: LI, Joachim J
APPLICANT: GAVIN, Kimberly
TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOEBACH, TEST, ALABRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,106
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 1676 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 13..1302
US-08-484-106-21

```

alignment_scores:

Quality: 105.00 Length: 278
Ratio: 0.802 Gaps: 13
Percent Similarity: 47.122 Percent Identity: 21.942

alignment_block:

US-09-578-063-47 x US-08-484-106-21 ..

Align seg 1/1 to: US-08-484-106-21 from: 1 to: 1676

```

85 ProLysLysThrGlySerArgProValValLeuLeuGlnHisGlyLeuVal 101
||||| ||||| ||||| ||||| |||||
85 CCAGAAAAAGAGAGATCCAGA.....CAGAAAAAGAC 116
101 IGlyGlyAlaSerAsnTrpIleSerAsnLeuProAsnAsn..... 114
117 GAATGGAAGAAAGATGCTTTCAGAAATTTGCATCAATTTAGAGAAG 166
115 .....SerLeuGlyPhe.....IleuAlaAspAla 123
167 ATTTGGAACAACTGGCTTCAGAGATGAAGTGTATCAATGGCTCAATCA 216
124 GlyPheAspValTrp.....MetGlyAsnSer.....Arg91 134
217 GCAATCGAAATTAATTAATGCAAGAAATGGGCTCAGACGAATGAA 266

```

```

134 yAenAlaTrpSerArgLysHisLys.....ThrLeu 145
    ||||| ||||| :
267 TAATGGGAATCCCTGCTGGAGAGCTGCTGAAATGGAATACCTGAG 316
    ||||| :
145 erlleAspGlnAspGlnPheTrpAlaPheSerTrpAspGluMetAlaArg 161
    ||||| :
317 AAATGAGGAGACGATGAGATCAGTAATGCTATCTACTGATTCACAAA 366
    ||||| :
162 pheAspLeuProAlaValIleAsnPhelIleGlnLysThrGlyGlnI 178
    ||||| :
367 TGTGATCTCCCTGGACTCGAATTAATTAACCAAAAAGATTAACGGA 416
    ||||| :
178 ulysIleTrpValGlyTyrSerGlnGlyThrThrMetGlyPheIle 195
    ||||| :
417 ATTCGAAAACGATTGAGCATCTCCGCGATTAATGATTTCGA..... 459
    ||||| :
195 lApheSerThrMetProGluLeuAlaGlnLysIleLysMetTrpPheAla 211
    ||||| :
460 .....AAATGGAAGCTTACCTAGCA 480
    ||||| :
212 LeuAlaProIleAlaThrValLysHisAlaLysSerProGlyThrLys 228
    ||||| :
481 GCTGGA..... 486
    ||||| :
228 eleuLeuLeuProAspMetMetIleLysGlyLeuPheGlyLysGluP 245
    ||||| :
487 .....TTTAATATCTTTTGGACGGGTGCTGCGTTCGACCGTCATG 526
    ||||| :
245 heLeuTrpGlnThrArgPheLeuArgGlnLeuValIleTrpLeuGly 261
    ||||| :
527 TTCTC.....ACGAATTGAGATGAGCTATCCGATTATACATATATG 570
    ||||| :
262 GlnValIleLeuAspGlnIleCysSerAsnIleMetLeuLeuGly 278
    ||||| :
571 AGAGTGGATGACGAGAAAGATGGCTCAATGTAAGTCTCTCTGAGC 620
    ||||| :
278 yPheAsnThrAsn.....AsnMetAsnMetSerArgAlaSerValT 292
    ||||| :
621 TATCAATGAGAAATATGAGCTGAATGTATGTGAAGAGAGC..... 663
    ||||| :
292 yAlaAlaHisThrLeuAlaGlyThrSerValGlnAsnIleLeuHisTrp 308
    ||||| :
664 .....CAATCTACGATTAGTTGG 681
    ||||| :
309 SerGlnAlaVal.....AsnSerGlyGluLeu..... 317
    ||||| :
682 GCTCGATCTATTGCGAGAAAATGAATAGCCACAGTTGATTCTTATCAT 731
    ||||| :
318 .....ArgAlaPheAspTrpGlySerGlu 325
    ||||| :
732 TGATATATTTGAGAGCTGCTGATGTGAAGAGTGAAT 765
    ||||| :
seq_name: /cgn2_6/ptodata/2/1na/5B_COMP.seq:US-08-602-359A-24
seq_documentation_block:
: Sequence 24, Application US/08602359A
: Patent No. 5942430
: GENERAL INFORMATION:
: APPLICANT: ROBERTSON, Daniel E.
: APPLICANT: MORPHY, Dennis
: APPLICANT: REID, John
: APPLICANT: MAFFIA, Anthony
: APPLICANT: LINK, Steven
: APPLICANT: SWANSON, Ronald V.
: APPLICANT: WARREN, Patrick V.
: APPLICANT: KOSMOTKA, Anna
: TITLE OF INVENTION: ESTERASES
: NUMBER OF SEQUENCES: 42
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FISH & RICHARDSON P. C.
: STREET: 4225 EXECUTIVE SQUARE, STE 1400
: CITY: LA JOLLA
: STATE: CALIFORNIA

```

```

: COUNTRY: USA
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 INCH DISKETTE
: COMPUTER: IBM PS/2
: OPERATING SYSTEM: MS-DOS
: SOFTWARE: WORD PERFECT 6.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/602,359A
: FILING DATE: February 16, 1996
: CLASSIFICATION: 435
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: HAILE, LISA A.
: REGISTRATION NUMBER: 38,347
: REFERENCE/DOCKET NUMBER: 09010/010001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-678-5070
: TELEFAX: 619-678-5099
: INFORMATION FOR SEQ ID NO: 24:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1041 NUCLEOTIDES
: TYPE: NUCLEIC ACID
: STRANDEDNESS: SINGLE
: TOPOLOGY: LINEAR
: MOLECULE TYPE: GENOMIC DNA
: US-08-602-359A-24

```

```

alignment_scores:
    Quality: 102.00      Length: 180
    Ratio: 1.052        Gaps: 6
    Percent Similarity: 53.889    Percent Identity: 22.778

```

```

alignment_block:
US-09-578-063-47 x US-08-602-359A-24 ..

```

```

Align seg 1/1 to: US-08-602-359A-24 from: 1 to: 1041

```

```

20 LeuIleuValAlaTrpMetPheGlnArgAsnValAsnSerValHisMe 36
    ||||| :
178 CTGCTCTGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 227
    ||||| :
36 tProth.....LysAlaValAspProGluAlaPheMetAsnI 49
    ||||| :
228 CCGGCTATGTGCGCTATGATGATGATGATGATGATGATGATGATGATG 277
    ||||| :
49 leSerGluIleIleGlnHisGlnGlyTrpProCysGluGluTrpGluVal 65
    ||||| :
278 TTGCGCGC.....GAGGGCTCACCTACCAAGAGCTTCAGCGTG 315
    ||||| :
66 AlaThrGlnAspGlyTrpIleLeuSerValAsnArgIleProArgGlyLe 82
    ||||| :
316 ACCGGGAGAGATGCTGTGTGTT.....CGGGCTG 347
    ||||| :
82 uValGlnProLysThrGlySerArgProValValLeuLeuGlnHisG 99
    ||||| :
348 GGTGCTGGGCCCGCGCGCTGGGGCAACCGGCTGCTTTGATGTGACG 397
    ||||| :
99 LysLeuValGly...GlyAlaSerAsnTrpIleSerAsnLeuProAsn 114
    ||||| :
398 GGTATACGTGGTCCGCTGCGCCGCCCTACATGCTGCTGCGCCGCGAG 447
    ||||| :
115 SerLeuGlyPheIleLeuAlaAspAlaGlyPheAspValTrpMetGly 131
    ||||| :
448 .....CTCGTGAAGTGGGGGTACCCGCTGCTGTGTGTCGA 482
    ||||| :
131 nSerArgGlyAsnAlaTrpSerArgLysHisLysThrLeuSerIleAsp 148
    ||||| :
483 CTTCGGGGCCACGGGAGAGCGGGGCTGACAGCAGATTGG..... 525

```

148 InAspGluPheTrpAlaPheSerTyrAspGluMetAlaArgPheAspLeu 164
 526 CCGCGGAGAGTGTGATGCC 546
 165 ProAlaValAlaIleAsnPheIleLeuGlnLysTrpGlyGlnGluLysIleTyr 181
 547 CCGGCGTGTGGTATGTCTCGAGAGCGGTCCCGCGCGGATMAAT 596
 181 rTyraValGlyTyrSerGlnGlyTrpThrMetGlyPheIle 194
 597 ATTGGTGGGTTACGATGAGCGCGCGCTGTACGATCGTG 636
 seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-232-519-1

seq_documentation_block:

Sequence 1, Application US/08232519
 Patent No. 5484725
 GENERAL INFORMATION:
 APPLICANT: Kageyama, Bunji
 APPLICANT: Nakae, Masanori
 APPLICANT: Yagi, Shigeo
 TITLE OF INVENTION: NORBORANE TYPE ESTER HYDROLASE
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Neave
 STREET: 1251 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10020
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/232,519
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 5-96286
 FILING DATE: 22-APR-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Haley Jr, James F.
 REGISTRATION NUMBER: 27,794
 REFERENCE/DOCKET NUMBER: SHGN-4
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 596-9000
 TELEFAX: (212) 596-9090
 TELEX: 14-8367
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1167 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 ORGANISM: Acetobacter pasteurianus
 STRAIN: ATCC 12873
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1164
 US-08-232-519-1

alignment_scores:

Quality: 98.50 Length: 400
 Ratio: 0.518 Gaps: 20
 Percent Similarity: 47.500 Percent Identity: 20.250

alignment_block:

US-09-578-063-47 x US-08-232-519-1

Align seg 1/1 to: US-08-232-519-1 from: 1 to: 1167

38 ThrLysAlaValAspProGluAlaPheMetAsnIleSerGluIleLeu 54
 250 ACCGCGATGTGATATGAGCCCTTATC 279
 54 nHsGlnGlyTyrProGlySerGluGluTyrGlu ValAlaThrGlu 69
 280 AAAAATTTCATGCAGAAAACCCACAGATGCGGTGCAGACGAG 325
 69 spGlyTyrIleLeuSerValAsnArgIleProArgGly LeuVal 83
 326 ATGCGCTGTGCGGGAGAACCCAGCAGGTATCGGTAGCGACACACGCTA 375
 84 GlnProLysLysThrGlySerArg ProValIleLeuGln 97
 376 AATGTGCGTGATGTTCGACGAGGCGACGCGCATTTGCTGTG... 423
 97 nHsGlyLeuValGlyGlyAlaSerAspTrpIleSerAsnLeuProAsn 114
 424 .CAGGTTTGGCGGAGATATGACGACACTGCTG... CTCACACAG 466
 114 snSerLeuGlyPheIleLeuAlaAspAlaGlyPheAspValTrpMetGly 130
 467 ATGCTTGCGCCGACAGAAAGCCGTAATTCCTTGTATCTG... 507
 131 AsnSerArgGlyAsnAlaTrpSerArgLysHis LysThrLe 144
 508 CCGGGCATGGGCTTCCTCTAATAACGTGGCAGACGCGCT 551
 144 uSerIleAspGlnAspGluPheTrpAlaPheSerTyrAspGluMetAla 161
 552 GCGGTTT 558
 161 rGpHeAspLeuProAlaValIleAsnPheIleLeuGlnLysTrpGlyGln 177
 559 TTGGCGGTGTGTGTAAGGAATGTGCAACCCCTTAATA 600
 178 GluLysIleTyrTyrValGlyTyrSerGlnGlyTrpThrMetGlyPheIle 194
 601 GAAAAAGCCCATGTGTGGCCATTTCTGGGG GGGGCGAT 641
 194 eAlaPheSerTrpMetProGluLeuAlaGlnLysIleLysMetTyrPhe 211
 642 TCCCCTGACCTGCTGCGAGATCACCTGATAGGTGGCACCTGAC 691
 211 LeuAlaProIleAlaThrValLysHisAlaLysSerProGlyThrLys 227
 692 TTTTGCCCCAGCCGGTTG GGTAAAGATGTGAATGCAGAT 732
 228 PheLeuLeuLeuProAspMetIleLysGlyLeuPheGlyLysGln 244
 733 TTTATC ACCGATTTGTGTGATAGTCA 758
 244 uPheLeuTyrGlnTrpArgPheLeuArgGlnLeuValIleTyr 258
 759 AAGCAGCCGCGATATGAAAGGCTGTTTGCATAATGCTGTATACAAAG 808
 259 .LeuCyGlyGlnValIleLeuAspGlnIleLysSerAsnIleMetLeu 274
 809 CCTATGTGGCGGTAAAGATGTGGATGCCGTG 840
 275 LeuLeuGlyGlyPheAsnThrAsnAsnMetAsnMetSerArgAlaSerVal 291
 841 CTGCGTGCACGTAGCGCTAGATGGCGCGCGGAT 873
 291 lTyraAlaAlaHisThrLeuAlaGlyThrSerValAlaAsnIleLeuHis 308
 874 CCGCTGCACGTTATGTCAAAGCGTCTCCCAAC 909
 308 rPserGlnAlaValAsnSerGlyGluLeuArgAlaPheAspTrpLysSer 324
 324

```
910 .....GGGCATCAGCGCGGATGATCTGCATCGGTGCTAGCTGGGCG 951
325 GluThrLysAsnLeuGluLysCysAsnGlnProThrProValArgTyrAr 341
|||||
952 GAACA..... 957
341 gValArgAspMetThrValProThrAlaMetTrpThrGlyGlnAspT 358
||||| :||| :|||
958 .....CCTACCCAGATTTTCTGGGGCCAGGAGATG 988
358 rPLeuSerAsnProGluAspValLysMetLeuLeuSerGluValThrAsn 374
::: :||| :|||
989 AAATTCTTCTGTCTCCACGCCCTGCGCCAGATGTCATCCCGTG 1038
375 LeuIleTyrHisLys.....AsnIleProGluTrpAlaHisValAspPh 389
::|||::|||::|||::|||
1039 ACAGTGTATGAAAGAACAGGCCCATCTGCCGACCTGAACATGCACAGA 1088
389 eIleTrpGlyLeuAspAlaProHisArgMetTyrAsnGluIleHisL 406
::: :||| :|||
1089 TGTG.....AACAAAGCCATGCCC 1108
406 eu..MetGlnGlnGluGluThrAsnLeuSerGlnGlyArgCysGluAla 421
|| :||| :||| :|||
1109 TGTTTGTAAGAACCCCGAAGCCGCGCTGAGCATGCCCCGGATGACGCG 1158
```



```

154 ..... 154
1303 TCGGAGAGACTGTTAGCAGATCCTGGGGATTCGACCATGTTCTGTTT 1352
155 ..... SerTyraSpG1 158
1353 GTTTATCATTTGCGAATACTTGCTCTTTCCCTTTGTGATGATGATGA 1402
158 uMetalaatgphaspLeuProAlaValIleasnPhelIleuGlnLysT 175
1403 AATGGCTAGGTTGACCTTCACAGTGAATGATTAATCTTACAGAAAA 1452
175 hnglynglnuLysIleTyrrTyValGlyrrSerGlnGlyrrThrMet 191
1453 CGGGCCGAAAAAGTCTATTATGCGGCTACACAGGCGCACCATCATG 1502
192 GlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLysIleYsme 208
1503 GCGTTATTGCAATTTCCACAAATGCCAGACTGACTCATATAAATCAAA 1552
208 tTyrrPheAlaLeuAlaProIleAlaThrValLysHsAlaLysSerPro 225
1553 GATTTTGCCTTAGCTCCCTATAGCCACTGTTAAATATGCAAGAGTCTG 1602
225 LyrThrLysPheLeuLeuLeuProAspMetMetIleLysGlyLeuPheGly 241
1603 GTACCAAAATTTCTGCTGCTCCAGATATGATGATGATATATATTTGGC 1652
242 LysLysGluPheLeuTyrrGlnThrArgPheLeuArgGlnLeuValIleTy 258
1653 AGCAAGAGTTTATATCCAGACTAGATTTTTCAGACAGCTTTTATTTA 1702
258 rLeuGlyGlyGlnValIleLeuaspGlnIleCysSerAsnIleMetLeuL 275
1703 CCTTGGGCCAGATGATCTTGAACCAATCTGACAGCAACATCATCTTAC 1752
275 euLeuGlyGlyPheAsnThrAsnAsnMetSerArgAlaSerVal 291
1753 TCCGGAGGTTTTTAACACAAATATGAAACAGACAGCAAAAGTGTG 1802
292 TyrAlaAlaIsthrLeuAlaGlyThrSerValGlnAsnIleLeuHsIst 308
1803 TAGTGCCCATACACTGCTGAGACTCTGTGAGAAATTCCTCCACTG 1852
308 pSerGlnAlaValAsnSerGlyLysLeuArgAlaPheAspTyrGlySerG 325
1853 GAGCCAGCGCGTGAATCTGGGAACTTCGCTTGACTGGGGAGTGTG 1902
325 LurhrLysAsnLeuGlnLysCysAsnGlnProThrProValArgTyrArg 341
1903 AGACAAAATAATCAGAGAAATGTAATCAGCAACTCTATAAGGTACAAA 1952
342 ValArgaspMetThrValProThrAlaMetThrGlyGlyGlnAspTyr 358
1953 GTTCGAGATTTGATGTCCTCAGACAGCAATGTGGCTGGAGTCAAGACTG 2002
358 pLeuSerAsnProGluAspValLysMetLeuLeuSerGlyValThrAsnL 375
2003 GCTTCAAAATCCAGATGATGTGAAAACATTAATCTTGAAGTAAACCA 2052
375 euIleTyrrHsLysAsnIleProGluThrPheAlaHsValAspPheIleTyr 391
2053 TCATCTACCAACAAGAACTTCGAGTGGCTCTGATGGATTCATCTGG 2102
392 GlyLeuaspAlaProHsIstArgMetTyrAsnGlnIleIleHsLysMetG1 408
2103 GGGGTGATGCCCCCTCAGCTGTTTACATGAATTTATACATCTCGAGAA 2152
408 ngLingLutThrAsnLeuSerGlnGlyArgCysGluAlaValLeu 423
2153 GCAG...GAGCCCAACCTCCCCAGGAGACCTCGAGGCTCAAAATTTG 2195

```

```

seq_name: gb_hlc:AK009571
seq_documentation_block:
LOCUS AK009571 1349 bp mRNA HTC 05-JUL-2001
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched
library, clone:2310032608, full insert sequence.
ACCESSION AK009571
VERSION AK009571.1 GI:12844447
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,
clone:11b:RIKEN full-length enriched mouse cDNA library
clone:2310032608.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.
JOURNAL High-efficiency full-length cDNA cloning
METHODS Methods in enzymology. 303, 19-44 (1999)
PUBMED 99279253
TITLE 10349636
AUTHORS 2 (bases 1 to 1349)
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.
Normalizaton and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
genome_research. 10 (10), 1617-1630 (2000)
JOURNAL 20459374
MEDLINE 11042159
PUBMED 11042159
REFERENCE
AUTHORS 3 (bases 1 to 1349)
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Kono,H., Akiyama,J., Nishi,K., Kikunishi,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
Iamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanabe,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A., and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multichannel sequencer
genome_research. 10 (11), 1757-1771 (2000)
JOURNAL 20530913
MEDLINE 11076861
PUBMED 11076861
REFERENCE
AUTHORS 4 (bases 1 to 1349)
The RIKEN genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5 (bases 1 to 1349)
ADACHI,J., AIZAWA,K., AKAHIRA,S., AKIMURA,T., AONO,H., ARAI,A.,
ARAKAWA,T., CARNINCI,P., FUKUDA,S., FUKUSHI,Y., FURUKO,M.,
HANEKAWA,T., HARA,A., HAYATSU,N., HIMOTO,K., HIROKA,T., HORI,F.,
IMOTOAN,K., ISHII,Y., ITOH,M., IZAWA,M., KATO,H., KAWAI,J.,
KOJIMA,Y., KONO,H., KOUDE,M., KOYA,S., KURIHARA,C., MATSUYAMA,T.,
MIYAZAKI,A., NISHI,K., NOMURA,K., NUMAZAKI,R., OHNO,M., OKAZAKI,Y.,
OKIDO,T., OWA,C., SAITO,H., SAITO,R., SAKAI,C., SAKAI,K., SANO,H.,
SASAKI,D., SHIBATA,K., SHIBATA,Y., SHINGAWA,A., SHIRAKI,T.,
SOGABE,Y., SUZUKI,H., TAGAMI,M., TAGAWA,A., TAKAHASHI,F.,
TANAKA,T., TEJIMA,Y., TOYA,T., YAMAMURA,T., YASUNISHI,Y.,
YOSHIDA,K., YOSHINO,M., MURAMATSU,M., and HAYASHIZAKI,Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@sc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to

```

Prepare more tissues. First strand cDNA was primed with a primer [5'-GAGACAGACGAGACGACCAAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGACAGAGTCTTCGAGCTTATTTAAATTATATCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

FEATURES

Source

Location/Qualifiers

```

source
1. 1349
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="MGD:MGI:1895763"
/db_xref="MGD:MGI:1914967"
/clone="2310032G08"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stge="adult"
52. 1239

```

BASE COUNT	396 a	290 c	284 g	379 t
ORIGIN				

```
alignment_scores:
  Quality: 1139.50
  Ratio: 3.485
  Percent Similarity: 79.562
  Length: 411
  Gaps: 4
  Percent Identity: 52.555
```

alignment_block:
US-09-578-063-47 x AK009571

Align seg 1/1 to: AK009571 from: 1 to: 1349

[illegible]

106	AsnPrp1IesSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAs	122
322	AACGGATTACAAATCTGCCAAACAGCCCTGGCGCTTCAATTCAGCA	371
122	palGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerA	139
372	TGCTGGCATATATGTGTGGCTGGGGACAGTCGAGGGAATACATGTGCC	421
139	IgLySHISLySthLeuSerIleAspGlnAspGluPheTrpAlaPheSer	155
422	GGAAAAATGTACTATTCCACAGACCTCAGTTGAATCTGGGGCTTCAGC	471
156	TyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLe	172
472	TTTATGGAAATGGCTAAATATAGACCTTCACACCACCATAGACTTCATGT	521
172	uGlnLySthrcLyGlnGlyuLySthIeTyrTyValGlyTrpSerGlnGlyT	189
522	ACAGAAAACTGGACAGAGGAAGATACACTATGTGGTCACTCCAGGCA	571
189	hTrpMetGlyPheIleAlaPheSerTrpMetProGluLeuAlaGlnLyS	205
572	CCACTATGCGTTTATTCCTCTTCTACCAATCCGCTCTGGCTAAAAA	621
206	IleuSmetYrPheAlaLeuAlaProIleAlaThrValLySHISAlaLy	222
622	ATCAAGAGGTTTTTATGTATGTAGCTCCACACTCTCCTAGTAGTATACAG	671
222	sSerProGlyThrLySthPheLeuLeuLeuProAspMetMetIleuSgLyL	239
672	AAGTCCTTTAAAAAGATTCACTTATTCCTAAGTTTCTCTCAAGGCGA	721
239	eupheGlyLySthGluPheLeu..... TyrGlnThrArgPheLeu	252
722	TATTTGGTACAAAAATGTCTTACGCCCAACACTATTGTGATCAATTTCTT	771
253	ArgGlnLeuValIleTyrLeuCySgLyGlnValIleLeuAspGlnIleCy	269
772	GGTACGGAGG..... TGCTCAGGGGAGCGTAGATCTCTCTG	812
269	sSerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnMetAsnM	286
813	CAGACAGCTTATATCACTCTCTGGGATTTGACAGAAAACTTAATATG	862
286	eTserArgAlaSerValTyrAlaIleAsnIleLeuAlaGlyThrSerVal	302
863	TGAAGTCGTTGATGTATCTAGGGCAATATCCAGCAGGAACTATCTACT	912
303	GlnAsnIleLeuHISTrpSerGlnAlaValAsnSerGlyGluLeuArgAl	319
913	CAAAACCTTTTCCACTGGGCAAGCTTCTAACTGGGAAGCTTCAAGC	962
319	alphaSPTPrGlySerGluThrLySthAsnLeuGluLyScysAsnGlnProT	336
963	CTAAACGGGGAGATCCATTACGAAATATGTTCACTACATACACGAAAA	1012
336	hTrpValArgTyrArgValArgAspMetTrpValProThrAlaMetTrp	352
1013	CGCTCCCTACTATGATGTGTGCACACCATGACCCGTCGCAATGTGCAGTGG	1062
353	ThrGlyGlnAspTrpLeuSerAsnProGluAspValLySmetLeuLe	369
1063	AAGCGGGCCATGACATCTGGCTGATCCCAAGATGCGCAATGTCTCT	1112
369	uSerGluValThrAsnLeuIleTyrHISLySthAsnIleProGluTrpAlaH	386
1113	TCCCAAACTCCCAACCTTCGTGTCCATAAAGAAATCTTCCATACATC	1162
386	ISValAspPheIleTrpGlyLeuAspAlaProHISArgMetTyrAsnGlu	402
1163	ACCTGGACTTCATTTGGGCGCATGGATGGCGCTCAAGAGGTTTACATAG	1212
403	IleIleHISLeuMetGlnGlnGluGlnThrAsn	413


```

1164 ACCTGACCTTACTCTGGCGATGATGCCCTCCAGAGCTTACATGAG 1213
403 I1e1le1s1leu1m1e1c1n1g1n1g1u1r1p1r1a1n 413
1214 ATAGTTACCATGATGATGCACAGACACTACAGAAAT 1246

seq_name: gb_hnc:AK010093
seq_documentation_block:
LOCUS AK010093 1350 bp mRNA HNC 05-JUN-2001
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched
library, clone:2310067K20, full insert sequence.
ACCESSION AK010093
VERSION AK010093.1 GI:12845297
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,
clone:2310067K20.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
1 (bases 1 to 1350)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Methods in enzymology. 303, 19-44 (1999)
99279253
PUBMED
10349636
2 (bases 1 to 1350)
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome research. 10 (10), 1617-1630 (2000)
20499374
PUBMED
11042159
3 (bases 1 to 1350)
Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsuki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome research. 10 (11), 1757-1771 (2000)
20530913
PUBMED
11076861
4 (bases 1 to 1350)
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5 (bases 1 to 1350)
Aachhi, V., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arikawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M.,
Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirakawa, T., Horii, F.,
Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J.,
Kojima, Y., Kono, H., Konda, M., Koye, S., Kurihara, C., Matsuyama, T.,
Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y.,
Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H.,
Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Tanaka, T., Teijima, Y., Toya, T., Yamamura, T., Yasunishi, A.,
Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (10-JUN-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
Riken Yokohama Developmental Biology Laboratory, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-resgsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

```

```

273 GCAGAGACCTGTCGATATTTGCACATGCTTATTCATCAGCCACA 322
106 AantPILSerAnLeuProAnSnSerLeuGlyPheIleuA1aAs 122
323 AACTGGATTACAATCTGCCAACAACACAGCTGGCTTCATCTACACA 372
122 p1aGlyPheAspValTTPMeGlyASnSerArgGlyASnAlaTTPSerA 139
373 TCCTGGCTATGATGTGTGGTGGGAACAGTCAGAGGAATACATGCTCC 422
139 rGLyHIS1s1rLeuSerIleAspGLAspGLAspGLAspGLAspGLAsp 155
423 GGAATAATATACATATTCACACAGCTGCTGCTGCTGCTGCTGCTGCT 472
156 TyRAspGLMeTAlaArgPheAspLeuProAlaValIleAsnPheIle 172
473 TTTGATGAATAGCTTAATATGACCTTCCAGCAGATGACTGATGCT 522
172 uG1n1yS1rG1yG1n1g1u1l1yS1l1e1y1r1y1a1g1y1r1S1e1r1G1n1g1y1T 189
523 ACAGAAACTGCACAGAGAGATACACTATGTGTGCTACTCTCAGGCA 572
189 hrThrMeG1yPheIleAlaPheSerThrMeProGLuLeuAlaGLn1yS 205
573 CCACATACGCTTATTCCTTCTTACCAATCTGCTGCTGCTGCTGCTGCT 622
206 I1e1yS1m1e1TyR1PheIleAlaLeuAlaProIleAlaThrVal1yS1H1s1Ala1y 222
623 ATCAAGAGGTTTATGATGATGATGATGATGATGATGATGATGATGAT 672
222 sSerProGLyThr1yS1rPheLeuLeuProAspMeT1l1e1yS1G1y1L 239
673 AAGTCCCTTAAAGATTTCATCTATTCCTAAGTTCTTCTTCAAGGTGA 722
239 euPheGL1yL1yS1rGLuPheLeu.....TyRG1n1r1h1r1rPheLeu 252
723 TTTTGGTACAAATGTTTCATGCCCCACACTACTTACATCAATTTCTT 772
253 ArgGLnLeuVal1l1e1TyR1euc1yS1GLn1Val1l1e1e1u1S1pGLn1l1e1cy 269
773 GGTACGGAAGT.....TGCTCAGGAGGCTGCTGCTGCTGCTGCTGCT 813
269 sSerAn1l1e1MeT1l1e1u1e1u1GLyGLyPheAsn1r1h1r1S1n1e1MetAsn1M 286
814 CAGCAACGCTTATTCATCTCTGTGATTCACAAAGAAACTTAAATG 863
286 e1SerArgAlaSerValTyR1AlaAlaH1s1ThrLeuAlaGL1yH1r1SerVal 302
864 TGAGTCGCTTATGATGTATCTAGGCAATATCCAGCAAGAACTACTACT 913
303 GLnAsn1l1e1u1H1s1r1SerGLn1AlaValAsnSerGLyGLuLeuArgAl 319
914 CAGACCTTTTCCACTGGGACAGCTGTCTAATCTGGGAAGCTTCAAGC 963
319 aPheAsp1r1rPGLySerGLuThr1yS1Asn1e1u1GLy1yS1yS1Asn1GLn1ProT 336
964 CTATACACGGGAGACTCCATTACGACACATGTTTACATCAATCAGAAAA 1013
336 hrProValArgTyRArgValArgAspMeTThrValProThralaMeT1rP 352
1014 CGCCTCCCTACTATGATGTGTCAGCCAGACCGTCCCAATGCTAGTGG 1063
353 ThrGLyGLyGLnAsp1r1rPLeuSerAnProGLuAspVal1yS1MeT1e1u 369
1064 AAGGTCGCAAGACATCTGCTGATCCCAAGATGTCGCAATCTCTCT 1113
369 uSerGLuValThrAsn1e1u1l1e1y1r1H1s1yS1Asn1l1e1ProGLu1r1PAlaH 386
1114 TCCCAACTCCCAACTTCTGTACATAGAGATCTTCCCAATC 1163
386 lSvalAsPheIle1rPGLyLeuAspAlaProH1s1ArgMeT1yRAsnGLu 402

```

COMMENT

please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGCAGAGAGAGATCCAGCAGCAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using triazole thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGCAGAGATTCGAGCTTATTTAAATATACCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI, 3' end: SstI. Host: SOLR.

FEATURES

Source	Location/Qualifiers
1.	1350
	/organism="Mus musculus"
	/strain="C57BL/6J"
	/db_xref="taxon:10090"
	/db_xref="MGI:1898392"
	/db_xref="MGI:1914967"
	/clone="231067K20"
	/sex="male"
	/tissue_type="tongue"
	/clone_lib="RIKEN full-length enriched mouse cDNA library"
	/dev_stage="adult"
	53. 1240

CDS

BASE COUNT
ORIGIN

396 a 290 c 286 g 378 t

AED

```
/dd_xref="GI:1284528"  
/translation="MLMLVTSVLSAAGAHGLGKLGPKNPKANMNVSONIWTYGVPR  
SESEYEVATEDGEDYGLIGVRIPIYQKKNNSRIPAYVLOHGHSAMATWITLIPNNLSLA  
FLIADAGYDWLGSNRSRGNTNRKSNKYVSPSEVMAQSPEDMAEDTPTADTFVQKTSR  
GOKRIHVHSGSGTIGTIFLAFSTNPDAKIKRYALAPAYVTSPEFKSILIRP  
FLIKVIEGNKEMPHNYLIDFQETEVASCRLDILGNALTFVFGGPKNNLNVSRPQV  
YLCHNPAGSTQDPLFHMAOLAKSGKLOAYNGSGSLQMWLHNLQKTPPYDASATFVPL  
AYVNGGHIDILADPODVAMLLPKLPNLILYKHELIPYINHLDFIWNADAQOEYVNEIYVMM
```

alignment_scores:

Quality:	1139.50	Length:	411
Ratio:	3.495	Gaps:	4
Percent Similarity:	79.319	Percent Identity:	52.798

alignment_block:

US-09-578-063-47 x AK010093

Align seg 1/1 to: AK010093 . from: 1 to: 1350

```

7   ArgILITRpleValSerIAsrMetGluMetRTPLeuLeuIleLeva 23
    |||::||| ::::: ::::: |||::: |||::: |||:::
35 AGGCACGTG.....CACATCAAGANGTGCGTCGATATGATAC 72
    :::::
23 IAlATyrmEtpeHgl..ArAsnVAlasnsErvalHisetProthrL 39
    :AAGTGTGTATCTGCACTTTGGAGGTGCCACATGCCCTATTGGAAACTG 122
73 ysaIaValaspProGluAlaphMeLasIIeSeGIlIleIGlnHis 55
    123 GTCCCAAAAACCCTGAAGCAAACATGAATGTTAGCAGATATPACTAC 172
56 GInGIyTyProcysGluLuTryrGlUValAlaThrGlsaspGlyTril 72
    |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
173 TGGGGATATCCAAGTAGGAATATGAAGTGTACTGTGAAGATGGCTCAT 222
72 elenSerValasnArgIleProrArgGlyLeuValGlnProIlysTrhg 89
    |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::

```

223	TCGGGGGCTATGAAATCTCTTAAGGGGAGAAAAATTCGAGTAATACG	272
89	LYSerATrProValValLeuLeuGlnHisGlyLeuValGlyValAser	105
111		111
273	GCMAAGACACCTGTGCATATTGTGCGAGCATGGTTTGTTGTCACCGACCA	322
106	AsnTrpIleSerAsnLeuProGlnSerLeuGlyPheIleLeuAlaAs	122
111		111
323	MACTGGATTACAAATCTGCGCAACACAGCCTGGCTTCATTCTTACGAGA	372
122	PAIaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAlaTrpSerA	139
373	TCCTGGCATATGATGTGTGGCTGGGAGACAGTCGAGGAATACATGTCTCC	422
139	rgLYsHisLYsThrLeuSerIleAspGlnAspGlnPheTrpAlaPheSer	155
111		111
423	GGAAAAAGTATACATTTCCACGACATCGATGGTAATCTGGCTTCAGCC	472
156	TrpAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIle	172
111		111
473	TTTTGATGAAAGGGCAATATATCCCTTCACGACCATATGACATTCATGTT	522
172	uGlnLYsThrGlyGlnGlyLYsIleEtyTrpValGlyTrpSerGlnLYT	189
111		111
523	ACGAGAAATCGACACAGAGAGATACATATGTGTGTGTCACCTCGAGCA	572
189	hTrpMetGlyPheIleAlaPheSerTrpMetProGlyLeuAlaGlnLYs	205
111		111
573	CCACATATCGTTTTATTTGGCTTTCTTACCAATCTCGCTGGCTGTAATAA	622
206	IleLYsMetTrpPheAlaLeuAlaLeuAlaTrpValTrpValLYsHisLYs	222
111		111
623	ATCAAGAGGTTTATATGACATATCTACCTCAAGTGTCTAGTGTGAATATACAGA	672
222	sSerProGlyTrpLYsPheLeuLeuLeuProAspMetIleLYsGlyL	239
111		111
673	AGTTCCTTTAAAAAGATTTCTCAATTTCCTTAAGTTTCTTCCAGAGTGA	722
239	eupheGlyLYsLYsLupheLeu.....TyrGlnTrhArpPheLeu	252
111		111
723	TATTGGTATACAAATATGTCATGGCCCACTACTATGATATTTCTT	772
253	ArgGlnLeuValIleEtyLeuGlysetGlyGlnValIleLeuAspGlnIleCy	269
111		111
773	GGTACGGAAGTG.....TGCACGGGGAGCTGATAGATCTTCTGTG	813
269	sSerAsnIleMetLeuLeuLeuGlyGlyPheAsnTrhAspMetAsnM	286
111		111
814	CAGCAACGCTTATATTCATATCTTGTGGATTTTACACAGAAAACTTAAATG	863
286	eSerArgAlaSerValTrpAlaAlaHisTrhLeuAlaGlyTrhSerVal	302
111		111
864	TAAGTTCGCTTATATGTGTATCTAAGGCAATATCCAGAGAACTATCTACT	913
303	GlnAsnIleLeuHisTrpSerGlnAlaValAlaAsnSerGlyGlnLeuArgAl	319
111		111
914	CAAGACCTTTTCCACATGGGACAGACTGTCTAATATGTGGAAAGCTTCAAGC	963
319	ArhAspTrpArgLYsSerGluTrhLYsAsnLeuGlyLYsCysAsnGlnProT	336
111		111
964	CTATATCTGGGGAGATCTTACAAACATATGGTACATCACTCAATCAAGAAA	1013
336	hTrpValArgTrpValArgValArgSerMetTrhValProTrhAlaMetTrp	352
111		111
1014	CGCTTCCTCCATATGATGTGTAGGCAACATGACCGTCCCATTTGCAATGTGG	1066
353	ThrGlyGlyGlnAlaTrpTrpLeuSerAsnProGlnAspValLYsMetLeuLe	369
111		111
1064	AACGGTGGGCAATGATCTGTGCTGATGTCGCCCAAGATGTGCAATGTCTGCT	1113
369	uSerGlyValTrhAsnLeuIleEtyHisLYsAsnIleProGluTrpAlaHis	386
111		111
1114	TCCCAACATCCCAACCTTCTGTATCCATTAAGAGATCTTCCCTCCATCAATG	1163

[illegible]

```

COMMENT
kanagawa 230-0045, Japan (E-mail:genome-res@gsic.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAAGACGATCCAGACGCTCTTTTCTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 5.0 and subtraction to
Rot = 25.0. Second strand cDNA was prepared with the primer adapter
of sequence [5'.
GAGAGAGACATTCTCAGGTAAATTAATTAATCCCCCCCCCCC 3']. cDNA was cleaved
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.

FEATURES
Source
Location/Qualifiers
1..1350
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="MGD:MGI:1898377"
/db_xref="MGI:MGI:1914967"
/clone="2310069P19"
/sex="male"
/tissue_type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
53..1240
/note="putative"
/codon_start=1
/protein_id="BAB26725.1"
/db_xref="GI:12845372"
/translacion="MWLLLYTVLSAFGAHGLFGRKPKPNKNANVYSONITYTFP
SEEEVTEDEGYILGVYRIPKRKNSNICKRPAYIVAQHGLIASATWITLPPNNSLIA
FIADAGDYWLGNRSRGNTSRKNVYSPOSPVEAFSSPEMAKYDDPATIDFVQRT
GOELIHVHGSOQTIGIFIASTPIALAKIKRYLAAPLATIVYTSPESFKSLIPK
FLKLVINGKMFMFNHYLDOLFETGEVCSRELLDLGSNALEIFCGFPKKNLNVRFDVP
YVNHPACTSTODLFHRAOLAKSGKIQAAYMWSPLONMLHNOKTPPYOVSATVPIII
AVNNGGDIIADPDVAMLLPKPLNLIYHKELIEPTNRDLPMADAPQEVYNELVTMM
AEV"

BASE COUNT      398 a      289 c      283 g      380 t
ORIGIN
Alignment_scores:
Quality: 1139.50          Length: 411
Ratio: 3.495              Gaps: 4
Percent Similarity: 79.319 Percent Identity: 52.798

alignment_block:
US-09-578-063-47 x AK010139 ..

Align seg 1/1 to: AK010139 from: 1 to: 1350

7 ARGGINTPRLIEVALSERHISARGMETGLUMETTRPLEULEULILEUVA 23
|||||            ::|||
35 AGGCACCTGG.....CACATCAAGAGATGGCTCTATTAGTAAC 72
|||||            ::|||
23 LALATYRMETPHGIN...ARGASNVALASNERVALHISMETPROTHL 39
::: |||         ::|||
73 AAGTGTCGATATCGATTGGAGGTGCACATGCGCTATTGGAAAACTGG 122
::|||           ::|||
123 GCCCAAACAACCCTGAGCAACATGAATGTAGTCAGATGATACTTAC 172
::|||           ::|||
56 GINGLYTYRRPGCYGLUGLUTYRGLIVALALATHRHGLUASGLYTYRIL 72
|||||            ::|||
173 TGGGATATCCAAAGTAGGAATATGAAGTGTACTGAAGATGGCTTACAT 222

```

72 eLeuSerValAsnArgIleProArgIleuValGlnProLysLysThrg 89
 223 TCTGGGGCTATAGAAATTCCTATGGGAGAAAATTCGAGAAATTCG 272
 89 LysArgProValIleuLeuGlnIleuValGlyGlyAlaSer 105
 273 GCAAGAGACCTGGCATATTCAGCATGTTGATTCATCAGCCACA 322
 106 AsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAs 122
 323 AACCTGGATTCACAAATCGCCAAACAGACCTGGCTTCATTCAGAGA 372
 122 PAlaGlyPheAspValTrpMetGlyAsnSerArgIleAsnAlaTrpSer 139
 373 TCGCTGGATGATGTGTGGCTGGGAAACAGTCGAGGAATTCATGCTCC 422
 139 rGlySHSLysTrpLeuSerIleAspGlnAspGluPheTrpAlaPheSer 155
 423 GGAATAATGTATCTATTCACACAGACTCAGTGAATTCCTGGCTTCAGC 472
 156 TyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIle 172
 473 TTGATGAATGCTGAATATGACTTCACGCCACATAGACTTCATGTGT 522
 172 uGlnLysTrpGlyGlnGluLysIleTyrTyrValGlyTyrSerGlnGlyT 189
 523 ACAGAAACAGGACCAAGAGAGATACACTAGTGTGCTCCTCAGGGA 572
 189 hTrpMetGlyPheIleAlaPheSerTrpMetProGluLeuAlaGlnLys 205
 573 CCACTATCGGCTTTATTCCTCTTCTACCAATCTGCTGCTGCTGCTGCT 622
 206 IleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysSHSLaLys 222
 623 ATCAAGAGGCTTTATGATTCATTCACCAATCTGCTGAGATACAGA 672
 222 sSerProGlyThrLysPheLeuLeuProAspMetIleLysGlyL 239
 673 AAGCCCTTTAAAGATTCATTCATTCATTCATTCATTCATTCATTCAT 722
 239 eUpheGlyLysLysGluPheLeu.....TyrGlnTrpArgPheLeu 252
 723 TATTGGTATCAAAATGTCTATGCCCACTACTGATGATCAATTCCTT 772
 253 ArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleCys 269
 773 GGTACGGAAGTG.....TGCTCAGGAGGCTGCTAGATCTCTCTG 813
 269 sSerAsnIleMetLeuLeuGlyGlyPheAsnTrpAsnMetLeu 286
 814 CAGCAAGCCTTATTCATTCCTGCTGATTCAGCAAGAAACCTAAATG 863
 286 eTserArgAlaSerValTyrAlaIleAsnTrpLeuAlaGlyThrSerVal 302
 864 TGAATTCCTTATGATGTATCTAGGCAATATCCAGCAAGCAATCTCT 913
 303 GlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGluLeuArgAl 319
 914 CAAACCTTTCCACTGGCAGACCTGCTAAATCTGGGAAGCTTCAAGC 963
 319 aPheAspTrpLysSerGluTrpLysAsnLeuGluLysCysAsnGlnProT 336
 964 CTATTAACGGGAGAGCTCATTACAGAACATCTTACACTACATCAGAAA 1013
 336 hProValArgTyrArgValArgAspMetThrValProThrAlaMetTrp 352
 1014 CGCTCCCTACTATGATGTATCAGCAAGACCGTGCCTCAATGCAAGTGG 1063
 353 ThrGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeu 369
 1064 AACGTGGCCATGACATCTGCTGATCCCAAGATGTCGCAATGCTGCT 1113

369 uSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAlaH 386
 1114 TCCCAATCCGCCCAACCTTTGTACCAATGAAGAGATTCCTCCCAACAT 1163
 386 tSValaAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnGlu 402
 1164 ACCTGACCTTCATCTGGCGCATGGATGCGCTCAAGAGGTTTACAAATG 1213
 403 IleIleHisLeuMetGlnGlnGluIleThrAsn 413
 1214 ATACTTACCATGATGCGCAAGAACTAACAGAAAT 1246
 seq_name: gb_hlc:AK009431
 seq_documentation block: 1352 bp mRNA HTC 05-JUL-2001
 LOCUS AK009431
 DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched
 library, clone:2310021B16, full insert sequence.
 ACCESSION AK009431
 VERSION AK009431.1 GI:12844222
 KEYWORDS CAP trapper.
 SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,
 clone:lib:RIKEN full-length enriched mouse cDNA library
 clone:2310021B16.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1352)
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome research. 10 (10), 1617-1630 (2000)
 JOURNAL 20499374
 MEDLINE 11042159
 REFERENCE 3 (bases 1 to 1352)
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuaki, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohata, E., Watanabe, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multichannel sequencer
 Genome research. 10 (11), 1757-1771 (2000)
 JOURNAL 20530913
 MEDLINE 11076861
 REFERENCE 4 (bases 1 to 1352)
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 TITILE Functional annotation of a full-length mouse cDNA collection
 JOURNAL Nature 409, 685-690 (2001)
 REFERENCE 5 (bases 1 to 1352)
 AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
 Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M.,
 Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirooka, T., Horii, F.,
 Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J.,
 Kojima, Y., Konno, H., Kouda, M., Koyu, S., Kurihara, C., Matsuyama, T.,
 Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, T.,
 Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H.,
 Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
 Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
 Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yanashiki, A.,
 Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission
 Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: <http://genome-gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT Please visit our web site (<http://genome-gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTAA 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence 15', GAGAGAGAGATCCAGAGTAAATTAATCCCCCCCCCC 3'. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOLR.

FEATURES

source

Location/Qualifiers

1..1352

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/db_xref="MGD:MGI:1894047"

/db_xref="MGD:MGI:1914967"

/clone="2310021B16"

/sex="male"

/tissue_type="tongue"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

55..1242

/note="putative"

/codon_start=1

/protein_id="BAB26283.1"

/db_xref="GI:12844223"

/translation="MMLLVTSVLSAFGAGLFGKLPKNPEANMNVSYMTWYR

SEEEVETEDYILGVYRIPIYKKNSENIKRPAVALOHGLIASATMTITLPPNNSLA

FLIADAGDVALISRGNTSRKNVYSPDSFEMWASFDMAKYDLPATIDIVQKT

GOEKIHYVHSGQTTIGFIASSTNPALAKTKRYPALAPVATYTESPFKISLIRK

VLGNVIFGNKMFHNYLDOFLGTEVCSRELIDLCNSALEIFCGFPKKNLNVSRLLV

FLGNPAGTSTODLPHMAQLAKSGKLQAYNMGSPLONNLHYNNQTPPYDYVSAMTVPI

AVMNGGHDILADPODVAMLLPKLPLNLYKELIPLYNHIDFMADADPOEYNEIVTM

AED"

BASE COUNT 397 a 291 c 285 g 379 t

ORIGIN

alignment_scores:

Quality: 1139.50 Length: 411
Ratio: 3.495 Gaps: 4
Percent Similarity: 79.319 Percent Identity: 52.555

alignment_block:

US-09-578-063-47 x AK009431 ..

Align seg 1/1 to: AK009431 from: 1 to: 1352

```

7  ATGGLTTPILLeValSerHisArgMetGluMetTPLeuLeuLeuVala 23
   |||:|||||
37 AGGCACTGG.....CACTACAGAGATGTGGCTGCTAATTAGTAAAC 74
   |||:|||||
23 lAlaTyMetPheGln..ArgAsnValaSerValHisMetProThL 39
   |||:|||||
75 AGGTGCTATGTCATTGGAGGTGCACATGCGCTAATTGGAAAACTGG 124
   |||:|||||
39 yAlaValaSprProGluAlaPheMetAsnHisSerGluLeuLeuGlnHis 55
   |||:|||||
125 GTCCCAAAAACCTGAGCAAAACATGAATGTAGTCAGATGATTAACACT 174

```

```

56 GInGlyTyProCysGluGluTyTrGluValAlaThGluAspGlyTrI 72
   |||:|||||
175 TGGGGATATCCAGTACGAGTAATGAAGTTGTACTGAAAGATGCTACAT 224
   |||:|||||
72 eLeuSerValAsnArgIleProArgGlyLeuValGlnProLysLysTh 89
   |||:|||||
225 TCTGGGGGCTATAGAAATTCCTATAGGAGAAAATTCGAAATATGTC 274
   |||:|||||
89 lYserArgProValAlaLeuLeuGlnHisGlyLeuValGlyGlyAlaSer 105
   |||:|||||
275 GCAAGAGACCTGTGCATATTTCCACATGCTTGAATGCAATGACACACA 324
   |||:|||||
106 AsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAla 122
   |||:|||||
325 AACTGATTAACAATCTGCCAAACACAGCTGCTTCTATCTACACAGA 374
   |||:|||||
122 pAlaGlyPheAspValTrpMetGlyLysSerArgGlyAsnAlaTrpSer 139
   |||:|||||
375 TGCTGGCTATGATGTGTGGGAGACAGTCCAGGGAATACATGGTCCC 424
   |||:|||||
139 rGlyHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSer 155
   |||:|||||
425 GGAATAATGTATCTATTCACCGACTCATGTAATTCGTGGCTTCACAG 474
   |||:|||||
156 TyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIle 172
   |||:|||||
475 TTTGATGAATATGCTAATATATGACTCCAGCCACCATAGACTTCATTG 524
   |||:|||||
172 uGlnLysThrGlyGlnGluLysIleTyrValGlyTyrSerGlnGly 189
   |||:|||||
525 ACAGAAACTGGACAGAGAGATACATATGTGTGCTCCTCCAGGGCA 574
   |||:|||||
189 hTrhMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLys 205
   |||:|||||
575 CCCTATCGCTTTATTCCTTTCTACCAATCTCTCTGCTTAATAAAA 624
   |||:|||||
206 IleLysMetTyrrPheAlaLeuAlaProIleAlaThrValLysHisAla 222
   |||:|||||
625 ATCAAGAGGTTTATCATTAAGTCCAGTGTGCTAGTGAATACAGAC 674
   |||:|||||
222 sSerProGlyThrLysPheLeuLeuLeuProAspMetIleLysGly 239
   |||:|||||
675 AAGTCCCTTAAAGATTCATTCCTTAAGTTCTCTCCAAAGTGA 724
   |||:|||||
239 euPheGlyLysLysGluPheLeu.....TyrGlnThrArgPheLeu 252
   |||:|||||
725 TATTGGTACAAATGTTCAATGCCCCACAACTACTTAATCAATTTCT 774
   |||:|||||
253 ArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIle 269
   |||:|||||
775 GGTACGGAAGTG.....TGCTCAGGAGAGCTGCTAGACTTCTCG 815
   |||:|||||
269 sSerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnAsnMetAsn 286
   |||:|||||
816 CAGCAACGCTTATTCATCTCTGTGGATTTCACAGAAAACCTTAAG 865
   |||:|||||
286 eSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVal 302
   |||:|||||
866 TGAGTGCCTTATGATGTCTAGGCAATATTCAGACAGAAACATCTACT 915
   |||:|||||
303 GlnAsnIleLeuHisTrpSerGlnAlaValaSerGlyGluLeuArgAl 319
   |||:|||||
916 CAAGACCTTTCACAGTGGACAGCTGTGAATATGCGGAACTTCAAC 965
   |||:|||||
319 aPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 336
   |||:|||||
966 CTATACTGGGAGAGTCCATTACAGAACATGTTAACATACTACGAAAA 1015
   |||:|||||
336 hProValArgTyrArgValArgAspMetThrValProThrAlaMetTrp 352
   |||:|||||
1016 CGCCTCCCTACATATGATGTGACAGCATACCGCTGCATTTCCAGTGTG 1065
   |||:|||||
353 ThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValLysMetLeu 369

```



```

1011 CGCCTCCCTACTATGATGTGTACGCCATGACCGCTGCCAATTGCAGTGTG 1060
353 ThrGlyGlyGlnAspTrrPleUserAsnProGluAspVallyMetLeuDe 369
AACGGTGGGCGATCATCGTGGCTGATGCCCAAGATGCGCATGTGCT 1110
1061 AAGGTGGGCGATCATCGTGGCTGATGCCCAAGATGCGCATGTGCT 1110
369 uSerGluValThrAsnLeuIleTyTrHisLysAsnIleProGluTrrAlaH 386
1111 TCCCAACACTCCCAACCTCTGTACCATTAAGGAGATTCCTCCCTACAATC 1160
386 IsValAspHeIleTrrGlyLeuAspAlaProHisArgMetTyrrAsnGlu 402
1161 ACCTGACCTTCATCTGGCGATGATGCGCCCTCCACAGGTTTACATGAG 1210
403 IleIleHisLeuMetGlnGlnGluGluThrAsn 413
1211 ATAGTTACCATGATGTCGACAGACATACAGAAAT 1243

seq_name: gb_htc:AK009473

seq_documentation_block:
LOCUS AK009473 1345 bp mRNA HNC 05-JUL-2001
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched
library, clone:2310022H17, full insert sequence.
ACCESSION AK009473
VERSION AK009473.1 GI:12844293
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,
clone:11b:RIKEN full-length enriched mouse cDNA library
clone:2310022H17.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.
1 (bases 1 to 1345)
TITLE High-efficiency full-length cDNA cloning
METHODS Methods in enzymology. 303, 19-44 (1999)
JOURNAL 99279253
MEDLINE 10349636
PUBMED 10349636
REFERENCE
AUTHORS 2 (bases 1 to 1345)
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.
1 (bases 1 to 1345)
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
METHODS Genome research. 10 (10), 1617-1630 (2000)
JOURNAL 20499374
MEDLINE 11042159
PUBMED 11042159
REFERENCE
AUTHORS 3 (bases 1 to 1345)
Shibata,K., Itoh,M., Alizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishii,K., Kitsuami,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamanoto,R., Matsunoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanishi,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsue,S., Kawai,D.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A., and Hayashizaki,Y.
1 (bases 1 to 1345)
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
METHODS Genome research. 10 (11), 1757-1771 (2000)
JOURNAL 20530913
MEDLINE 11076861
PUBMED 11076861
REFERENCE
AUTHORS 4 (bases 1 to 1345)
The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium
TITLE Functional annotation of a full-length mouse cDNA collection
METHODS Nature 409, 685-690 (2001)
JOURNAL 5 (bases 1 to 1345)
DEFINITION
AAchawa,T., Alizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
Hasegawa,T., Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M.,
Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hirooka,T., Horii,F.,
Imotoani,K., Ishii,Y., Itoh,M., Izawa,M., Kato,H., Kawai,J.,
Kojima,Y., Konno,H., Konda,M., Koya,S., Kurihara,C., Matsuyama,T.,
Miyazaki,A., Nishii,K., Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y.,
Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H.,

```


23	lAlAtyMeRpheIn..ArgaAnvaAlAsnSerValHtsMetProThL	39
24	
73	AAAGTGCCTATTCTGCATTGGAGGTGCACAAGGCCCTATTGGAAAACTGG	122
74	
39	ysAlaValAspProGlnAlaPheMetAsnIleSerGluIleIleGlnHis	55
40	
123	GTCCCAAAAAACCCGGAACCAACATGAAATGTTAGTCAGATGATTAAC	172
41	
56	GlnGlyTyPProCysGluGluTyTyrGluValAlaThrGluAspGlyTyrI	72
42	
173	TGGGGAATATCCAACTGAGGAAATATGAAGTACTGTTACTGAAGATGGCTACAT	222
43	
72	eLeuSerValAsnArgIleProAlaGlyLeuValGlnProLysThng	89
44	
223	TCGTGGGGCTCTATGAAATTCCTATTGGGAGAGAAAAATCTGGAAATATGC	272
45	
89	lySerArgProValValLeuLeuGlnHisGlyLeuValGlyValAser	105
46	
273	GCAAGAAACCTGTGGCATATTGGACGATGGTTGATTTCATTCAGCCACA	322
47	
106	AsnThrIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAs	122
48	
323	AACGCGATTCAAAATCTGCCAAACAACAGCCGGCCTCATCTTACACAGA	372
49	
122	PaLaGlyPheAspValTyrMetGlyAsnSerArgGlyAsnAlaTyrSerA	139
50	
373	TGCGGCTCATCATGTGTGGCTGGGGAACAGTCGAGGGAATACATGGTCCC	422
51	
139	rglySHLysThrLeuSerIleAspGlnAspGluPheThrPalaPheSer	155
52	
423	GGAAAAATGTAFACTATTACACAGACTCAGTTGAAATCTGGGGCTTTCAGC	472
53	
156	TyrAspGluMetAlaArgPheAspPheProAlaValIleAsnPheIleLe	172
54	
473	TTTGATGAAATGGCTAAATATGACCTTCCAGCCACACATAGACTTCATGT	522
55	
172	uGlnLysThrGlyGlnGluLysIleTyrTyTyrValGlyTyrSerGlnGlyT	189
56	
523	ACAAGAAACCTGCACAGAGAAAGATACACTATGTGGTCACTCCAGGGCA	572
57	
189	hThrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLys	205
58	
573	CCACTATCGGTTTAAATCCCTTTTACCAATCCCTGCTGGCTAAAAAA	622
59	
206	IleLysMetTyrPheAlaIleuAlaProIleAlaThrValLysHisAlaTy	222
60	
623	ATCAAGAGGTTTATGCACTTACGCTCAGTGTCTACTGTGAAGATACAGA	672
61	
222	sSerProGlyThrLysPheLeuLeuLeuProAspMetIleLysGlyL	239
62	
673	AAAGTCCCTTAAAAAAGAAATTCACCTTATTCCTAAGTTCTCTCAAGGTGA	722
63	
339	eupheGlyLysLysGluPheLeu.....TyrGlnThrArgPheLeu	252
64	
723	TATTTGTGAACAAAATGTTCAAGCCCAACACTACTAGATCAATTTCTT	772
65	
253	ArgGlnLeuValIleTyrLeuLysGlyGlnValIleLeuAspGlnIleCy	265
66	
773	GGTACGGGAAGTG.....TGCTCACGGGAGAGCTCTGATCTTCTCTG	813
67	
269	sSerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnAsnMetAsnM	286
68	
814	CAGCAAGCCTTATTCATCTCTGTGGAGATTGACAGAAGAAACCTTAATG	863
69	
286	eSerArgAlaSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVal	302
70	
864	TGAGTCCCTTTGATGTGATCTAGGGCAATAACTCACAGACAGAACATCTACT	913
71	
303	GlnAsnIleLeuHisThrSerSerIleValAsnSerGlyLysLeuArgAl	319
72	
314	CAGAGCTTTTCCACTGGCCACAGGCTGCTTAATCTGGGAGAGCTTCAAGC	963
73	

seq_name: gb_htc:AK009729

seq_documentation_block:

LOCUS AK009729 1346 bp mRNA HTC 05-JUL-2001
DEFINITION Mus musculus adult male tongue cDNA, full-length enriched library, clone:z310040L03, full insert sequence.

ACCESSION AK009729 GI:12844702
VERSION AK009729.1
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library
clone:z310040L03.

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Artibeus; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
METHODS Methods in enzymology. 303, 19-44 (1995)
PUBMED 99279253
MEDLINE 10349636

REFERENCE AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome research. 10 (10), 1617-1630 (2000)

JOURNAL MEDLINE 20499374
PUBMED 11042159

REFERENCE AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kitunai,T., Tashiro,H., Itoh,M., Suni,I., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kasaiwaqi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Metabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome research. 10 (11), 1757-1771 (2000)

JOURNAL MEDLINE 20530913
PUBMED 11076861

REFERENCE AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5 (bases 1 to 1346)
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,

Align seg 1/1 to: AK010116 from: 1 to: 1347

```

7 ArgGlnTrpIleValSerHisArgMetGluMetTrpIleuLeuIleuVal 23
  |||.....|
33 AGCGATGCG.....CACTACAGATGCGCTGCTATTATGATGATAC 70
23 AlaTyrMetPheGln...ArgAsnValAsnSerValHisMetProThrL 39
  |||.....|
71 AAGTGTCTATCTGCATTGTGGAGTGCACATGGCTATTGGGAACATCG 120
39 ysaIaValAspProGluAlaPheMetAsnIleSerGluIleIleGlnHis 55
  |||.....|
121 GTCCCAAAACCCGTGAAGCAACATGATGTAGTCAGATGATCAACTTAC 170
56 GlnGlyTrpProGlyGluTyrGluValAlaIleThrLysProGlyTyrI 72
  |||.....|
171 TGGGGATATCCAACTGAGTGAATGAAATGTTACTGAAATGGCTACAT 220
72 eIeSerValAsnArgIleProArgIleValAlaGlnProLysIleThrG 89
  |||.....|
221 TCTGGGGGTATATGATTCCTTATGGGAAGAAATTCGAGAAATTCG 270
89 IySerArgProValIleuLeuGlnHisGlyLeuValGlyGlyAlaSer 105
  |||.....|
271 GCAAGAGACCTGTGCATATTTCAGCATGTTGATTGCATCAGCCACA 320
106 AsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleuAlaAs 122
  |||.....|
321 AACTGGATTCAAATCTGCCAAACACAGCCCTTCATTCATTCAGAGA 370
122 PalaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAlaTrpSer 139
  |||.....|
371 TCGTGGCTATGATGTGTGCTGGGAACACTCGAGGAATACATGAGGCC 420
139 rGlyHisIlyStrIleuSerIleAspGlnAspGluPheTrpAlaPheSer 155
  |||.....|
421 GGAATAATGATATCTATTACACAGACTCAGTGAATCTGGGCTTACGC 470
156 TyrAspGlnMetAlaArgPheAspLeuProAlaValIleAsnIleIle 172
  |||.....|
471 TTGATGTAATGGCTAAATATGACCTTCACACACATACATTCATCTAT 520
172 uGlnIlyStrGlyGlnGluIlyStrIleTyrTyrValGlyTyrSerGln 189
  |||.....|
521 ACGAAACACGACAGACAGAGATACACTGTGTTGGTCACTCAGAGGCA 570
189 hrThrMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLys 205
  |||.....|
571 CCACATACGGTTTATTTGCTTTTACCAATCTGCTGCGCTAAATAAAA 620
206 IleLysMetIlyrPheAlaLeuAlaProIleAlaThrValIlyshIle 222
  |||.....|
621 ATCAAGAGGTTTATGATTAAGCTCCAGTGTCTACTGTAAGATATACAG 670
222 sSerProGlyThrLysPheLeuLeuLeuProAspMetIleLysGlyL 239
  |||.....|
671 AATGCCCTTAAAGATTTCACTTATCTTAAGTTCTTCTCAAGGTGA 720
239 euPheGlyLysGluPheLeu.....TyrGlnThrArgPheLeu 252
  |||.....|
721 TATTGTGAACAAATGTTTCATGCCCAACACTACTAGATCAATTTCT 770
253 ArgGlnLeuValIleTyrLeuGlyGlnValIleLeuAspGlnIleCys 269
  |||.....|
771 GGATCGAAGTG.....TGCCTCAGGAGCTCTAGATCTTCTCTG 811
269 sSerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnAsnMetAsn 286
  |||.....|
812 CACAAACGCTTATTCATCTCTGTGGATTTGACAGAAACCTTAAAG 861
286 eIeSerArgIleSerValTyrAlaAlaHisThrLeuAlaGlyThrSerVal 302
  |||.....|

```

```

862 TGAATCCCTTGTGTCTATCTAGAGCATATCCAGCAGGACATCTACT 911
303 GlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyLeuAlaG 319
  |||.....|
912 CAGACCTTTTCCACTGGGACACGCTTGTAAATCTGGGAAGCTTCAGC 961
319 aPheAspTrpGlySerGluThrLysAsnLeuGluLysCysAsnGlnPro 336
  |||.....|
962 CTATTAATCGGGAGATTCATTACAGAACATGTTTACTTAATTCAGAAA 1011
336 hrProValArgTyrArgValArgAspMetThrValProThrAlaMet 352
  |||.....|
1012 CGCTCCCTACTATGATGTGTGACCATGACCGTCCCAATTCAGATGTGG 1061
353 ThrGlyGlyGlnAspTrpLeuSerAsnProGluAspValIlyshIle 369
  |||.....|
1062 AACGGTGGCCATGACATCTGGCTGAGACCCCAAGATCTCCAAATGCG 1111
369 uSerGluValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAla 386
  |||.....|
1112 TCCCAAACTCCCAACCTTCTGTACCATAGAGGATTCCTCCATACATC 1161
386 IyValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnG 402
  |||.....|
1162 ACCTGACCTCATCTGGGCGATGGATGGCGCTCAAGAGGTTTCAATGAG 1211
403 IleIleHisLeuMetGlnGlnGluThrAsn 413
  |||.....|
1212 ATAGTTACCATGATGGCAGAGACATCAACAGAAAT 1244
seq_name: gb_hnc:AK010058
seq_documentation_block:
LOCUS AK010058 1348 bp mRNA HTC 05-JUL-2001
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched
library, clone:2110066C03, full insert sequence.
ACCESSION AK010058
VERSION AK010058.1 GI:12845238
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain: C57BL/6J) adult male tongue cDNA to mRNA,
clone:2110066C03.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1348)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Methods in enzymology. 303, 19-44 (1999)
99279253
MEDLINE
PUBMED
10349636
2 (bases 1 to 1348)
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome research. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
PUBMED
11042159
3 (bases 1 to 1348)
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsuno, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system - 384-format
sequencing pipeline with 384 multicapillary sequencer
Genome research. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
PUBMED
11076861
4 (bases 1 to 1348)

```


Quality: 1138.50 Length: 411
 Ratio: 3.492 Gaps: 4
 Percent Similarity: 79.319 Percent Identity: 52.555

alignment_block:

US-09-578-063-47 x AK009300

Align seg 1/1 to: AK009300 from: 1 to: 1349

```

7  ArgGlnTrrPleValSerHisArgMetGluMetTrpLeuLeuVala 23
   |||:|||||
35  AGGCACTGG.....CACTACAAGATGTGGCTGCTAATAGTAAC 72
   |||:|||||
23  LalaTrrMetPheGln...ArgAsnValAsnSerValHisMetProHrl 39
   |||:|||||
73  AAGTGTGCTATCTGCATTTGGAGGTGCACATGGCTATTTGGAAACTGG 122
   |||:|||||
39  ysaIaValAspProGluAlaPheMetAsnIleSerGluIleIleGlnHis 55
   |||:|||||
123  GTCCCAAAACCTGAAAGCAAAACATGATGTAGTCAGATGATCTTAC 172
   |||:|||||
56  GlnGlyTrrProCysGluTrrGluValAlaThrGluAspGlyTrrI 72
   |||:|||||
173  TGGGATATTCAGATGAGAAATGAAATGTAGTCTGAAGATGGCTACAT 222
   |||:|||||
72  eleuSerValAsnArgIleProArgGlyLeuValGlnProLysLysThrG 89
   |||:|||||
223  TCTGGGGGTATATAGAAATCTTATGGGAGAAATAATTCAGAAATATCG 272
   |||:|||||
89  LysArgTrpProValValLeuLeuGlnHisGlyLeuValGlyValAlaSer 105
   |||:|||||
273  GCAGAGACCTGTGGCATTTTGCAGCATGTGTTGATTCATCGACCCACA 322
   |||:|||||
106  AsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuAlaAs 122
   |||:|||||
323  AACTGATTCACAATCTGCCAAACACAGCTGCGCTTCATCTACAGAGA 372
   |||:|||||
122  palagIyPheAspValTrrMetGlyAsnSerArgGlyAsnAlaTrrPera 139
   |||:|||||
373  TGTGTGCTATGATGTGTGGCTGGGGAACACTCGGGGAATACATGCTCCC 422
   |||:|||||
139  rglYshLysLysThrLeuSerIleAspGlnAspGluPheTrrPalaPheSer 155
   |||:|||||
423  GGAAGAAATGATATCTATTCACACAGCACTGATGATCTGGGCTTACG 472
   |||:|||||
156  TyrAspGluMetAlaArgPheAspLeuProAlaValIleAsnPheIleLe 172
   |||:|||||
473  TTTGATGAAGATGGCTAATATGACCTTCACGCCACATGACCTCATTTGT 522
   |||:|||||
172  uGlnLysThrGlnGlnLysLysIleTrrYrrValGlyTrrSerGlnLysT 189
   |||:|||||
523  ACAGAAACAGTGAAGAGAGATACACTATGTGGTGCATCTCAGGSCA 572
   |||:|||||
189  hrhMetGlyPheIleAlaPheSerThrMetProGluLeuAlaGlnLys 205
   |||:|||||
573  CGACTATCGGTTTATTTGCTTTCTACCAATCTGCTGCTGGCTAAAAA 622
   |||:|||||
206  IleLysMetTrrPheAlaLeuAlaProIleAlaThrValLysHisAlaLys 222
   |||:|||||
623  ATCAAGAGGTTTATGATGATTAGCTCCAGTGTGCTAGTGAAGTATACAGA 672
   |||:|||||
222  sSerProGlyThrLysPheLeuLeuLeuProAspMetMetIleLysGlyL 239
   |||:|||||
673  AACTCCCTTTAAAAAGATTTCACTTATTCCTAAGTTCTCTCTCAAGGTGA 722
   |||:|||||
239  euPheGlyLysLysGlnPheLeu.....TrrGlnTrrArgPheLeu 252
   |||:|||||
723  TATTGTGTAACAAATGCTCATGCCCCACACAATCTAGATCAATTTCTTT 772
   |||:|||||
253  ArgGlnLeuValIleTrrLeuCysGlyGlnValIleLeuAspGlnIleCys 269
   |||:|||||
773  GGTACGGAATGTG.....TGCTCACGGAGAGCTGCTATCATCTTCTCTG 813

```

```

269  sSerAsnIleMetLeuLeuLeuGlyGlyPheAsnThrAsnAsnMetAsnM 286
   |||:|||||
814  CAGCAAGCGCTTATTCATCTCTGTGTGATTTGACAGAAATACTAATAG 863
   |||:|||||
286  eSerArgAlaSerValTrrAlaAlaHisThrLeuAlaGlyThrSerVal 302
   |||:|||||
864  TGAGTCCGCTTTGGTGTATCTAGAGGCATATCCAGCAGACATCTACT 913
   |||:|||||
303  GlnAsnIleLeuHisTrrPserGlnAlaValAsnSerGlyLysLeuAla 319
   |||:|||||
914  CAAGACCTTTTCCACGTGGGACAGCTTGTAAATCTGGGAAGCTTCAAG 963
   |||:|||||
319  aPheAspTrrGlySerGlnThrLysAsnLeuGluLysCysAsnGlnProT 336
   |||:|||||
964  CTATTAATGGGGAGAGTCCATTAACAGACATGTTACACTAATACAGAAA 1013
   |||:|||||
336  hrProValArgTrrArgValArgAspMetThrValProHrlAlaMetTrr 352
   |||:|||||
1014  CGCCTCCCTACTATGATGTGTGACCATGACCGCTGCAATTCAGTGTGG 1063
   |||:|||||
353  ThrGlyGlyGlnAspTrrPleuSerAsnProGluAspValLysMetLeu 369
   |||:|||||
1064  AAGGTTGGCCATGATCCTCGCTGATCCCAAGATGTCGCAATGCTGCT 1113
   |||:|||||
369  uSerGlnValThrAsnLeuIleTrrHisLysAsnIleProGluTrrPalaH 386
   |||:|||||
1114  TCCCAAACTCCCAACCTCTGTGTACCTAAGAGATTTCTCCCTACATC 1163
   |||:|||||
386  IsValAspPheIleTrrPglLeuAspAlaProHisArgMetTrrAsnGlu 402
   |||:|||||
1164  ACCGGAATTCATCTGGCGCATGATGATCGCTCAAGAGTTCATCAATGAG 1213
   |||:|||||
403  IleIleHisLeuMetGlnGlnGluGluThrAsn 413
   |||:|||||
1214  ATAGTTACATGATGAGCAGAACTACAGAAAT 1246
   |||:|||||
seq_name: gb_hlc:AK009474

seq_documentation_block:
LOCUS AK009474 1349 bp mRNA HTC 05-JUL-2001
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched
library, clone:2310022103, full insert sequence.
ACCESSION AK009474
VERSION AK009474.1 GI:12844295
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,
clone.lib:RIKEN full-length enriched mouse cDNA library
clone:2310022103.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1349)
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
METHODS In enzymology. 303, 19-44 (1999)
PUBMED 99279253
JOURNAL Methods in enzymology. 303, 19-44 (1999)
PUBMED 10349636
REFERENCE 2 (bases 1 to 1349)
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome research. 10 (10), 1617-1630 (2000)
PUBMED 20499374
PUBMED 11042159
REFERENCE 3 (bases 1 to 1349)
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

```


BASE COUNT 396 a 290 c 284 g 379 t
ORIGIN

alignment_scores:

Quality: 1138.50 Length: 411
Ratio: 3.492 Gaps: 4
Percent Similarity: 79.319 Percent Identity: 52.555

alignment_block:

US-09-578-063-47 x AK009546 ..

Align seg 1/1 to: AK009546 from: 1 to: 1349

```

7  ArgGlnTrpIleValSerHisArgMetGlnMetTrpLeuLeuVala 23
   |||||
   |||||
   |||||
38 AGGCACTGG.....CACTACAAAGATGGCTGCTATATAGTAC 75
23 LAlaTyrMetPheGln...ArgAsnValAsnSerValHisMetProThrL 39
   |||||
   |||||
   |||||
76 AAGTGTGCTATCTGCATTGGAGTGCACATGGCTATTGGAAAACTGG 125
39 ysaIaValAspProGlnIaPheMetAsnIleSerGlnIleIleGlnHis 55
   |||||
   |||||
   |||||
126 GTCCCAAAACCCTGACCAACATGATGTAGTCAGATGATACCTTAC 175
56 GlnGlyTyrProCysGlnGluTyrGlnValaIaThrGlnAspGlyTyrI 72
   |||||
   |||||
   |||||
176 TGGGATATCCAGTGAAGATATGAGTTGTACTGAAGATGGCTACAT 225
72 eleuSerValAsnArgIleProArgGlyLeuValGlnProLysLysThrG 89
   |||||
   |||||
   |||||
226 TCTGGGGCTATAGATCTTAAATCGTGAAGAAAAATCTGAGATATACG 275
89 lySerArgProValaIleLeuLeuGlnHisGlyLeuValaGlyLysIleSer 105
   |||||
   |||||
   |||||
276 GCAAGAGACCTGTGTCATATTGTCAGCATGGTTTATGTCATCACCACA 325
106 AsnTrpIleSerAsnLeuProAsnAsnSerLeuGlyPheIleLeuVala 122
   |||||
   |||||
   |||||
326 AACTGATTAACAATCTGCCAACAACAGCTGGCTTCATCTGACAGA 375
122 pAlaGlyPheAspValTrpMetGlyAsnSerArgGlyAsnAlaTrpSera 139
   |||||
   |||||
   |||||
376 TGTGGCTATGATGTGTGGTGGGACAGTCGAGGGAATACATGCTCC 425
139 rGlyHisLysThrLeuSerIleAspGlnAspGluPheTrpAlaPheSer 155
   |||||
   |||||
   |||||
426 GGAATAATGTACTATTACACAGCTCAGTTGATCTGGGCTTCCAGC 475
156 TyrAspGlnMetAlaArgPheAspLeuProAlaValIleAsnPheIle 172
   |||||
   |||||
   |||||
476 TTGTGTAATGGCTAATAATGACCTTCCAGCCACCATAGACTTCAATGT 525
172 uGlnLysThrGlyGlnGluLysIleTyrTyrValGlyTyrSerGlnGly 189
   |||||
   |||||
   |||||
526 ACAGAAACTGGACAAAGAGATACACTATGTGTGTCACCTCAGGGCA 575
189 hrThMetGlyPheIleAlaPheSerThrMetProGlnLeuAlaGlnLys 205
   |||||
   |||||
   |||||
576 CCACTATCGGTTTATTCCTTTCTTACCAATCCGCTCGCTGCTGAAAAA 625
206 IleLysMetTyrPheAlaLeuAlaProIleAlaThrValLysHisAla 222
   |||||
   |||||
   |||||
626 ATCAAGAGCTTTATGATAGCTCAGCTTCTACTGTAAGTATACAGA 675
222 sSerProGlyThrLysPheLeuLeuProAspMetIleLysGlyL 239
   |||||
   |||||
   |||||
676 AAGTCCCTTAAAAAGATTTCATTATCTCAAGTTCTTCTCAAGGTGA 725
239 euPheGlyLysLysGluPheLeu.....TyrGlnThrArgPheLeu 252
   |||||
   |||||
   |||||

```

```

726 TATTGGTAACAAATGTTCAAGCCCACTACTATGATCAATTTCTT 775
253 ArgGlnLeuValIleTyrLeuCysGlyGlnValIleLeuAspGlnIleC 269
   |||||
   |||||
   |||||
776 GGTACGGAAGTG.....TGCTACGGAGAGCTGTGATCTTCTGTG 816
269 sSerAsnIleMetLeuLeuGlyGlyPheAsnThrAsnAsnMetAsnM 286
   |||||
   |||||
   |||||
817 CAGCAAGCCTTATTCATCTCTGTGAGATTGACACAGAAAACTTAATG 866
286 etSerArgAlaSerValTyrAlaIaHisThrLeuAlaGlyThrSerVal 302
   |||||
   |||||
   |||||
867 TGATGCGCTTGATGTGATCTAGAGCATTAATCCAGAGAACTTACT 916
303 GlnAsnIleLeuHisTrpSerGlnAlaValAsnSerGlyGlnLeuArgAl 319
   |||||
   |||||
   |||||
917 CAAGACCTTTCCACTGGGACAGAGCTGCTAAATCTGGGAAGCTTCAAGC 966
319 aPheAspTrpGlySerGluThrLysAsnLeuGlnLysCysAsnGlnProT 336
   |||||
   |||||
   |||||
967 CTATTAAGTGGGAGAGTCCATTACAGACATGTTACACTACAAATCGAATA 1016
336 hrProValArgTyrArgValaArgAspMetThrValProThrAlaMetTrp 352
   |||||
   |||||
   |||||
1017 CGCCTCCCTACTATGATGTGTCAAGCATGACCGTCCCAATGCAAGTGTGG 1066
353 ThrGlyGlyGlnAspTrpLeuSerAsnProGlnAspValLysMetLeu 369
   |||||
   |||||
   |||||
1067 AACGGTGGCCATGACATCTGCTGATCCCAAGATGTGCGCAATGCTGCT 1116
369 uSerGlnValThrAsnLeuIleTyrHisLysAsnIleProGluTrpAlaH 386
   |||||
   |||||
   |||||
1117 TCCCAAACTCCCAACTCTGTGTACCAAGAGAGATTCTTCCCAATTC 1166
386 lSValAspPheIleTrpGlyLeuAspAlaProHisArgMetTyrAsnGlu 402
   |||||
   |||||
   |||||
1167 ACCTGAGCTTCACTGGGCGATGAGTGGCCTCAAGAGGTTTACATGAG 1216
403 IleIleHisLeuMetGlnGlnGluGluThrAsn 413
   |||||
   |||||
   |||||
1217 ATAGTTACCATGATGCGAAGACTAACAGAAAT 1249

```


Gile 199

Seem

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2002, 22:09:19 ; Search time 67.96 Seconds
(Without alignments) 461.051 Million cell updates/sec

Title: US-09-578-063-47
2247

Perfect score: 1 MERTSRQWVSHRHEMLL.....IHLMOEFTNLSGRCFAVL 423
Sequence:

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /SIDSR/gcgdata/geneseq/AA1980.DAT:*
2: /SIDSR/gcgdata/geneseq/AA1981.DAT:*
3: /SIDSR/gcgdata/geneseq/AA1982.DAT:*
4: /SIDSR/gcgdata/geneseq/AA1983.DAT:*
5: /SIDSR/gcgdata/geneseq/AA1984.DAT:*
6: /SIDSR/gcgdata/geneseq/AA1985.DAT:*
7: /SIDSR/gcgdata/geneseq/AA1986.DAT:*
8: /SIDSR/gcgdata/geneseq/AA1987.DAT:*
9: /SIDSR/gcgdata/geneseq/AA1988.DAT:*
10: /SIDSR/gcgdata/geneseq/AA1989.DAT:*
11: /SIDSR/gcgdata/geneseq/AA1990.DAT:*
12: /SIDSR/gcgdata/geneseq/AA1991.DAT:*
13: /SIDSR/gcgdata/geneseq/AA1992.DAT:*
14: /SIDSR/gcgdata/geneseq/AA1993.DAT:*
15: /SIDSR/gcgdata/geneseq/AA1994.DAT:*
16: /SIDSR/gcgdata/geneseq/AA1995.DAT:*
17: /SIDSR/gcgdata/geneseq/AA1996.DAT:*
18: /SIDSR/gcgdata/geneseq/AA1997.DAT:*
19: /SIDSR/gcgdata/geneseq/AA1998.DAT:*
20: /SIDSR/gcgdata/geneseq/AA1999.DAT:*
21: /SIDSR/gcgdata/geneseq/AA2000.DAT:*
22: /SIDSR/gcgdata/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2247	100.0	423	22	Human TANGO 294.
2	2076	92.4	390	22	Human TANGO 294 ma
3	1289	57.4	399	22	Human shear stress
4	1289	57.4	399	22	Human lysosomal ac
5	1166	51.9	398	14	RGL precursor Or
6	1166	51.9	221	22	Human TANGO 294 ex
7	1156.5	51.5	395	6	AAF50322
8	1141.5	50.8	379	15	AAF56870
9	1141.5	50.8	379	17	AAF50382
10	1141.5	50.8	380	15	AAF56871
11	1119	49.8	398	7	AAF60724

12	1119	49.8	398	7	AAF60658	Sequence of human
13	1119	49.8	398	17	AAW09383	Human gastric lipa
14	1119	49.8	398	22	AAF66086	Human lipase prote
15	783	34.8	144	22	AAF66070	Human TANGO 294 cy
16	691	30.8	233	22	AAF61608	Human protein HP03
17	514.5	22.9	410	21	AAV32309	Soybean acid triacyl
18	502	22.3	410	21	AAV32307	Rice acid triacyl
19	465	20.7	405	21	AAV32310	Soybean acid triacyl
20	398	17.7	286	21	AAV32304	Corn acid triacyl
21	262	11.7	157	21	AAV32303	Catalpa acid triacyl
22	171	7.6	33	22	AAF66066	Human TANGO 294 si
23	160.5	7.1	125	21	AAV32305	Corn acid triacyl
24	150	6.7	143	21	AAV32301	Corn acid triacyl
25	138	6.1	90	21	AAV32306	Corn acid triacyl
26	138	6.1	116	21	AAV32302	Catalpa acid triacyl
27	128	5.7	25	22	AAF66069	Human TANGO 294 tr
28	105	4.7	430	18	AAW22234	C. elegans origin
29	105	4.7	430	18	AAW14140	Caenorhabditis ele
30	102	4.5	346	18	AAW30688	Pyrodicticum sp. es
31	101	4.5	308	17	AAW31905	Pseudomonas wiscon
32	100.5	4.5	286	17	AAW31903	Pseudomonas wiscon
33	100.5	4.5	286	18	AAW22309	P. wisconsinensis m
34	99	4.4	678	22	AAW32271	C glutamicum prote
35	98.5	4.4	388	15	AAW31939	Acetobacter pasteu
36	98	4.4	471	20	AAW00878	Sucrose fatty acid
37	98	4.4	806	20	AAV21797	Alteromonas lipase
38	98	4.4	809	20	AAV21796	Alteromonas lipase
39	96.5	4.3	1010	21	AAW38012	Human secreted pro
40	95	4.2	336	22	AAW39139	Human polypeptide
41	94.5	4.2	340	21	AAW30361	Arabidopsis thalia
42	94.5	4.2	392	21	AAW20360	Arabidopsis thalia
43	94.5	4.2	1280	21	AAW4943	Human secreted pro
44	93.5	4.2	484	21	AAW45186	Arabidopsis thalia
45	93	4.1	637	22	AAW90967	C glutamicum prote

ALIGNMENTS

RESULT 1	AAW6065	standard; Protein: 423 AA.
AC	AAW6065:	
DT	30-MAR-2001 (first entry)	
XX	Human TANGO 294.	
XX	TANGO protein: INTERCEPT protein; neurological disorder;	
KW	central nervous system; focal brain disorder; bipolar affective disorder;	
KW	global diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;	
KW	senile dementia; Huntington's disease; amyotrophic lateral sclerosis;	
KW	Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;	
KW	neuropsychiatric; psychoactive substance use; anxiety.	
OS	Homo sapiens.	
XX	WO200077239-A2.	
PN	21-DEC-2000.	
XX	24-MAY-2000; 2000WO-US14858.	
PF	14-JUN-1999; 99US-0333159.	
XX	(MILL-) MILLENNIUM PHARM INC.	
PA	McCarthy SA, Fraser CC, Sharp JD, Barnes TM;	
PI	WPI; 2001-032313/04.	
XX	N-PSDB; AAF45131, AAF45132.	
DR		
XX		

1

Db 361 daphrmynelhlmgdeetnlsqrcceavl 390

RESULT 3

AAB90783 standard; Protein: 399 AA.

AC AAB90783;

DT 15-JUN-2001 (first entry)

DE Human shear stress-response protein SEQ ID NO: 66.

KW Human; shear stress-response protein; vascular disease;

KM arteriosclerosis.

OS Homo sapiens.

PN MO200125427-A1.

PD 12-APR-2001.

PF 02-OCT-2000; 2000WO-JP06840.

PR 01-OCT-1999; 99JP-0280976.

PA (KYOW) KYOMA HAKKO KOGYO KK.

PI (NOJIMA) NOJIMA H.

PI Nojima H, Yoshisue H, Obayashi M, Ota T, Kawabata A, Sakurada K;

PI Kuga T, Sekine S, Nakamura Y, Sugano S;

DR WPI: 2001-266308/27.

DR N-PSDB; AAH02906.

PT DNA sequences, proteins encoded by them and antibodies against them

PT useful in diagnosis and treatment of vascular disease caused by

PT arteriosclerosis -

PS Claim 60; Page 402-404; 678pp; Japanese.

CC The present invention provides the protein and coding sequences of a

CC number of human shear stress response proteins. These are useful in the

CC diagnosis, treatment and screening of vascular diseases caused by

CC arteriosclerosis, including heart failure, post-PTCA restenosis and

CC hypertension.

CC Sequence 399 AA;

Query Match 57.4%; Score 1289; DB 22; Length 399;

Best Local Similarity 63.5%; Pred. No. 1.5e-122;

Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

40 AVDPAPFAMISITLHOGCPCEYEATVDGTYILSNRIPRLVQPKKGRPVYLLDHG 99

DB 28 avdpapfamisitlhogcpceyevatedgtyilcnrphrkhsdkgkprvvlqng 87

100 LVGASNMNINSLPNNISLGLADAGFDVWGNMGRNAGNRKKTSLISDDEFAFSDM 159

DB 88 lvgasnmninslppnnislgldagfdvwmgnstrgtwsktkltsvsgdefafsydgm 147

160 AFEDLPVAVINFLQKTGQEKIYVGSQGTMGFIASFSTMPLEAOKIKMYPALAPIATVK 219

DB 148 afedlpvavinflktgqekiyvgsqgtmgfiastmpelakrikmffafgvasva 207

220 HAKSGTFFLLPDMKIRGLRKEKELVQFRLRQVLYLCOGYLLDQCSNMLLLGCF 279

DB 208 fctspmakirgripdhlkldlgdkeflpsatlkwtgltvchvllkelcgnclflgcf 267

280 NTNNNMNRASVYAAHTAGTGVONILHWSQAVNSGELRAPDMGSETKMLKCNQDTPPR 339

DB 268 ntnnmnrasyvyaahtagtgvonilhwsqavnsgeleapdmgsstklmlecnqdprr 327

QY 340 YRVDFVTPTAMTGGQDLSPEDYKMLSEVNLITAKNIPENAHYDFIWLDPARM 399

DB 328 yrvdfvtptamtgqdlspedykmlsevnlitaknipenahydfiwlclaparm 387

QY 400 YNEIIHMQO 409

DB 388 ynklihmrc 397

RESULT 4

AAB66061 standard; Protein: 399 AA.

AC AAB66061;

DT 30-MAR-2001 (first entry)

DE Human lysosomal acid lipase protein.

KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;

KW central nervous system; focal brain disorder; bipolar affective disorder;

KW global diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;

KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;

KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;

KW neuropsychiatric; psychoactive substance use; anxiety.

OS Homo sapiens.

PN MO200077239-A2.

PD 21-DEC-2000.

PF 24-MAY-2000; 2000WO-US14858.

PR 14-JUN-1999; 99US-0333159.

PA (MILL-) MILLENNIUM PHARM INC.

PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

DR WPI: 2001-032313/04.

PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for

PT screening assays and diagnostic assays and for the treatment of

PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's

PT disease -

PS Claim 8; Fig 6; 359pp; English.

CC The present invention relates to TANGO or INTERCEPT proteins and coding

CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,

CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding

CC sequences are useful for the treatment of neurological disorders such as

CC central nervous system (CNS) disorders, CNS-related disorders, focal

CC brain disorders, global diffuse cerebral disorders and other

CC neurological and cerebrovascular disorders. The CNS disorders include

CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic

CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,

CC autonomic function disorders such as hypertension and sleep disorders,

CC neuropsychiatric disorders, psychoactive substance use disorders,

CC anxiety, and bipolar affective disorder. The present sequence is a

CC sequence used in a sequence homology alignment with the TANGO/INTERCEPT

CC sequences of the present invention.

SO Sequence 399 AA;

Query Match 57.4%; Score 1289; DB 22; Length 399;

Best Local Similarity 63.5%; Pred. No. 1.5e-122;

Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

40 AVDPAPFAMISITLHOGCPCEYEATVDGTYILSNRIPRLVQPKKGRPVYLLDHG 99

PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -
 XX
 PS Claim 8; Page 326; 359pp; English.
 XX
 CC The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF5121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS diseases include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.
 XX
 SQ Sequence 221 AA;

Query Match 51.8%; Score 1165; DB 22; Length 221;
 Best Local Similarity 100.0%; Pred. No. 2.5e-110;
 Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 VHMPTKADPFAFMNISIIQHOGYPCFEYEVATEDGYILSYNRIPRGIVQKKTGSRV 93
 Db 1 vhmptkavdpfaafmnsiilqhgypcfeeyevatedgyilsvnrprglvqpkktgsrpv 60
 QY 94 VLOHGLVGASNMVSNLPPNNSLGLADAGFDVMMGNSRGNAMSRKHKTLSIDODEFWA 153
 Db 61 vllqhgylvgasnmvsnlppnnsilgladagfdvmmgnsrgnawsrkhkltlsiddqefwa 120
 QY 154 FSYDEMARFDLPAYINFLQKTGQEKIYYVGSQGTMGFIASFMPBELAQRIKMFALA 213
 Db 121 fsydemarfdlpayinfilqktgqekiyvgsqgtmgfiastmpbelagrikmfala.180
 QY 214 PIATYKHAKSPEKTKLLPDMAMIKGLFGKKEFLYQTRFLRQ 254
 Db 181 piatvkhaksptkelllpdmimkglfgkkellyqtrflrq 221

RESULT 7

AAFS0322
 ID AAF50322 standard; protein; 395 AA.

XX
 AC AAF50322;

XX
 DT 17-JAN-1992 (first entry)

XX
 DE Rat lingual lipase protein.

XX
 KW Lingual lipase; enzyme; EC-3.1.1.3; ss.

XX
 OS Rattus rattus.

XX
 PN GB2142337-A.

XX
 PD 16-JAN-1985.

XX
 PF 29-JUN-1984; 84GB-0016581.

XX
 PR 01-JUL-1983; 83GB-0017989.

XX
 PR 05-SEP-1983; 83GB-0023759.

XX
 PA (CELL-) CELLTech LTD.

XX
 PI Carey NH, Williamson R;

XX
 DR WPI; 1985-014450/03.

XX
 DR N-PSDB; AAN50385.

PT New lingual lipase protein for treatment of lipase deficiency - also

PT new pre-lingual lipase protein and related products
 XX
 PS Disclosure; Fig 5; 15pp; English.
 XX
 CC This protein may be expressed in a transformant host organism
 CC and may be used for the treatment of lipase deficiency.
 XX

SQ Sequence 395 AA;

Query Match 51.5%; Score 1156.5; DB 6; Length 395;
 Best Local Similarity 54.7%; Pred. No. 4.5e-109;
 Matches 216; Conservative 69; Mismatches 109; Indels 1; Gaps 1;

QY 17 MMLLLVAVMFO-RVNVSVHMTKAVDEPAFMNISIIQHOGYPCFEYEVATEDGYILSY 75
 Db 1 mmlillvavistfggahglfgkl9pnpaamnnisqmtlywypcfeeyevatedgyil9v 60
 QY 76 NRIPRGIVQKKTGSRPVLLQHGLVGASNMVSNLPPNNSLGLADAGFDVMMGNSRGN 135
 Db 61 yrtpngknsenlqkrprvyylqhgllaaatnlanlpnnsiafmladagfdvmlnsrgn 120
 QY 136 AMSRKHKTLSIDODEFWAFSYDEMARFDLPAYINFLQKTGQEKIYYVGSQGTMGFI 195
 Db 121 tswrknyvypdsveafsfidemakypatinfivqktgqekihygsqgttlgfia 180
 QY 196 FSTMBELAQRIKMFALAIPATVKAHKSPTGTFLLPDMAMIKGLFGKKEFLYQTRFLRQ 255
 Db 181 fstmbelakktltyalpvatlvkysplkksiflptflfklmtygkkmflphtyfdfl 240
 QY 256 VYLGOVYLDQICSNIMLLGFTNNMNMNRASVYAAHTLAGTSVONILHMSQAVNSG 315
 Db 241 gvevcsrevldlcsntlflfcgfdknlvnsrfdvlyghmpagcsvqdlwaglvrsq 300
 QY 316 ELRARDMGSEETNLEKCNQPRFVRVVRDMVTPAMTGGQDWLSNPEDEVKLLSEVTNL 375
 Db 301 kfgatfwgspqgmlyhqktrpeydvsaemtyvavawngndlladpqdvamllpklsnl 360
 QY 376 IYHKNIPENAHDFITGLDAPHRMNEYIHLMOE 410
 Db 361 lfhekellaynhldifwamdapevynemlsmaad 395

RESULT 8

AAFS6870
 ID AAR56870 standard; Protein; 379 AA.

XX
 AC AAR56870;

XX
 DT 20-FEB-1995 (first entry)

XX
 DE Canine gastric lipase.

XX
 KW Canine gastric lipase; CGL; dog; fat; mucoviscidosis; enzyme;
 KW bioconversion; exocrine pancreatic insufficiency.

XX
 OS Canis familiaris.

XX
 PN WO9413816-A.

XX
 PD 23-JUN-1994.

XX
 PF 16-DEC-1993; 93WO-FR01260.

XX
 PR 16-DEC-1992; 92FR-0015201.

XX
 PA (LJOU) INST RECH JOUVEINAL.

XX
 PI Benicourt C, Blanchard C, Junten J;

XX
 DR WPI; 1994-217890/26.

XX
 DR N-PSDB; AAO68388.

ID	AA	Score	DB	Length	380
AC	AAAS6871	50.8%	141.5	15	380
AC	AAAS6871	56.5%	Pred. No. 1.4e-107		
DF	20-FEB-1995	Conservative	67	Mismatches	90
XX		Indels	5	Gaps	3
DE	Canine gastric lipase (Met-CGL).				
XX	Canine gastric lipase; CGL; dog; fat; mucoviscidosis; enzyme;				
KW	bioconversion; exocrine pancreatic insufficiency.				
XX	Canis familiaris.				
XX	MO9413816-A.				
PN	23-JUN-1994.				
PD	16-DEC-1993;	93MO-FR01260.			
XX	16-DEC-1993;	92FR-0015201.			
PR	16-DEC-1992;	92FR-0015201.			
XX	(LJOU) INST RECH JOUVEINAL.				
PI	Benicourt C, Blanchard C, Junien J;				
XX	WPI: 1994-217890/26.				
DR	N-PSDB; AA068389.				
XX	Recombinant canine gastric lipase and nucleic acid encoding it -				
PT	are used for improving absorption of ingested fat, treating				
PT	mucoviscidosis etc. and in enzymatic bio-conversions				
XX	Claim 14; Fig 9A; 52pp; French.				
CC	The sequence given below is the sequence of figure 9A, altered				
CC	according to the amendments described on page 2 of the appended				
CC	letter.				
CC	CGL is used to improve absorption of ingested fat, in healthy and				
CC	sick patients (e.g. having altered levels of gastric lipase); to				
CC	treat conditions associated with insufficiency (or lack) of lipases,				
CC	esp. mucoviscidosis or exocrine pancreatic insufficiency and partic.				
CC	where immobilised, for bioconversions, e.g. hydrolysis or				
CC	transesterification (other mammalian gastric lipases, or derivs.,				
CC	can be used in this application).				
XX	Sequence	380 AA;			
XX	Query Match	50.8%	Score 141.5;	DB 15;	Length 380;
XX	Best Local Similarity	56.5%;	Pred. No. 1.4e-107;		
XX	Matches	210;	Conservative	67;	Mismatches 90;
XX				Indels	5;
XX				Gaps	3;
QY	37 PPKAADPEAFNMISELHOGQPCPEYEATDEGYLSVNRIPRLGYOKKTKGSRPVLL	96			
DB	8 PL---NPEVNMISQMLTYWGYPAEYEVTEQGYLIGDIRPYGKNSENIGRPVAFI	64			
QY	97 OHGLVGASNMWISLNPNNSLGFIILADAGFVMMGNSRGNAWSRKHTLSIDODEFAFESY	156			
DB	65 qhgllasatnwslpnpslatelladagdvwlgnsrgrtwarmllypsdevefawfsf	124			
QY	157 DEMARFDPAYVNFILQTKGQEKIYVGSQGTWGFIAFSFMPELAQIKKIFALAPIA	216			
DB	125 demakxydipatlidifllkktgqdklhvysggticiglaftstnpklakrikftyalapa	184			
QY	221 TYKHAHSPETKFLLLPDMWIKGLFGKKEFLYQTRFLRO-LTVIYLCGOVILIDICSNIML	275			
DB	185 lkytyecellinklmlypsstfllkllfgnklf-yphnifdqlatavdesretvdnlcnaali	243			
QY	276 LGGFITNNNNASRASYAAHTLACTAGSVONILHMSQAVNSGELRAFDGSGSETNLEKNOF	335			
DB	244 lgdfidmnlmsrldyishnpagcsyqvnvlwsgvksqkfgatdwsqpyqnmhnyhs	303			
QY	336 TPVRYRVARDMTVPTAMWIGGQDMLSNPEDEVKMLLSEVTNLIIHKNIPEMAHVDFTWGLDA	395			

```

Db      304 mppyyntldcmhyplavwngndlladphndvdlleklnillyhrklyppynhlidfwanda 363
Oy      396 PHRYATNEITILHM 407
        | :||||| : :|
Db      364 pqavyneivnmm 375
        | :||||| : :|

RESULT 11
AAP60724
ID AAP60724 standard; Protein: 398 AA.
AC AAP60724;
DT 08-JUN-1991 (first entry)
DE Sequence of pregastric lipase.
KW Precursor polypeptide; secretion vector; enzyme.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..19
FT /label= signal
FT Protein 20..398
PN MO8603778-A.
PD 03-JUL-1986.
PF 23-DEC-1985; 85WO-GB00599.
PI 21-DEC-1984; 84GB-0032483.
PR 23-DEC-1985; 85MO-GB00599.
PR 01-JAN-1986; 86GB-0019568.
XX (BREM/) BREWM RES FOUND.
PA (TUBB/) TUBB R S.
PI Tubd RS;
XX WPT: 1986-182910/28.
DR N-PSDB; AAN60685.
CC The inventors claim a DNA sequence encoding AAP60723 linked to a
CC promoter upstream and a gene for a polypeptide downstream.
CC Particular examples are the yeast enzyme AMG, the mammalian enzyme,
CC gastric lipase and the mammalian lymphokine, interferon-alpha2.
XX SQ Sequence 398 AA;

Query Match 49.8%; Score 1119; DB 7; Length 398;
Best Local Similarity 52.2%; Pred. No. 3e-105;
Matches 210; Conservative 73; Mismatches 107; Indels 12; Gaps

Oy 17 MWLLILVAFMFQRRNNSVMPTKAV-----DPEAFNNISRIIOHOGYPCEEYEVED 69
    ||||| : : : | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 mwllilvtmsall-----svlgtghlgfklhpspevtmnsqmtlywygneeyevted 55

Oy 70 GTILSVNRIPKGLVOPPKKTGRSPVYLQHGLVGASNMWISNLPPNSSLGFTLADAGFDVYM 129
    ||||| ||||| | | ||||| ||||| | | ||||| ||||| ||||| ||||| |||||
Db 56 gyllewnrlpygkxnsgrtqgryvfqlpgyllasatwnslpnnsflilladagdydw1 115

Oy 130 GNSRGNAMSRKKKITSLIDODEFMARFSDPMARFDLPAYINFLTKTGOKIYYGVXSOGT 189
    ||||||| |:| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 116 gasrignwtarlllyspdsvelwtatsfdemakydldpacitdlfvkkrtqgdqnlmyvghsqgt 175

```

```

OY 190 TMGFIATSTMPBLAKIKMYFALAPIATVKAHAKSPGTFKLLPDMKIGLFGKKEFLYOT 249
DB 176 tlgfiatstmpslakritfyalapvatvkytkslinklrvpsglfkfifgdklfyphn 235
OY 250 RFLROLVYLYLGOVILDOICSNIMLLGGFNTNNMNSRASVYAHAHTLAGTSVONILHMS 309
DB 236 ffdqfllatevcsremnlilcsnalfilcgfskfnftrldvylshnpagtsvqnmfhw 295
OY 310 QAVNSGELRAFDWSETEKNEKCNQPTPVRYRVDMYVPTAMWTGGODMLSNPEDVKML 369
DB 296 qavksqkfaydwsqpyvgrmhdydqgppyyvntamvplavwngskdlldpqqvyl 355
OY 370 SEVTNLTYHKNIPEMAHVDFTWGDAPHRMNEIHLMOOE 411
DB 356 pklpnllyhkeipfynhldfiwmdapevyndivsmiseck 397

RESULT 12
AAP60658
ID AAP60658 standard; Protein; 398 AA.
XX
AC AAP60658;
XX
DT 22-AUG-1991 (first entry)
XX
DE Sequence of human pregastric lipase.
XX
KW Cystic fibrosis therapy; enzyme; lipase deficiency.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT Protein 20..398
XX
PN WO601532-A.
XX
PD 13-MAR-1986.
XX
PF 15-AUG-1985; 85WO-GB00364.
XX
PR 21-AUG-1984; 84GB-0021210.
PR 15-AUG-1985; 85MO-GB00364.
PR 01-JAN-1986; 86GB-0008957.
XX
PA (CELL-) CELLTech LTD.
PA (LOWE/) LOWE P A.
XX
PI Lowe PA;
XX
DR WPI; 1986-081634/12.
DR N-PSDB; AAN60566.
XX
PT New gastric lipase protein, esp. of human origin - for treating
PT lipase deficiency, and DNA sequences coding for it
XX
PS Disclosure; Fig 3; 39pp; English.
XX
CC The inventors claim a pregastric lipase protein and a gene encoding
CC it. Gastric lipase protein is useful for oral administration to
CC treat lipase deficiency, e.g. cystic fibrosis or pancreatitis.
XX
SQ Sequence 398 AA;

Query Match 49.8%; Score 1119; DB 7; Length 398;
Best Local Similarity 52.2%; Pred. No. 3e-105;
Matches 210; Conservative 73; Mismatches 107; Indels 12; Gaps 2;
OY 17 MWLLILVAYMFORNVSNHPTKAV-----DPEAFMISEIIOHGRCCEYEYATD 69
DB 1 mwllilmasl-----svlgthglfgklhnpgspevmnlsgmltywypneeyevtd 55

```

```

OY 70 GYILSVNRIIPRGVLQPKTGSRPVLLQHLVGASAMNISLNPNNISGLFIADAGFDVWM 129
DB 56 gylilsvnrilprgvlqpkktgsrppvllqhlvgasamnislnpnnisglfiadagfdvwm 115
OY 130 GNSGNMNSRKHKLSIDODEFNAFSYDEMARFPLPAINIFILKGTGEEKIYYVYSGCT 189
DB 116 gnsrgnltwarrnllyspdsvefwaftidemakypalidivkvktgqkqihyvhsgct 175
OY 190 TMGFIATSTMPBLAKIKMYFALAPIATVKAHAKSPGTFKLLPDMKIGLFGKKEFLYOT 249
DB 176 tlgfiatstmpslakritfyalapvatvkytkslinklrvpsglfkfifgdklfyphn 235
OY 250 RFLROLVYLYLGOVILDOICSNIMLLGGFNTNNMNSRASVYAHAHTLAGTSVONILHMS 309
DB 236 ffdqfllatevcsremnlilcsnalfilcgfskfnftrldvylshnpagtsvqnmfhw 295
OY 310 QAVNSGELRAFDWSETEKNEKCNQPTPVRYRVDMYVPTAMWTGGODMLSNPEDVKML 369
DB 296 qavksqkfaydwsqpyvgrmhdydqgppyyvntamvplavwngskdlldpqqvyl 355
OY 370 SEVTNLTYHKNIPEMAHVDFTWGDAPHRMNEIHLMOOE 411
DB 356 pklpnllyhkeipfynhldfiwmdapevyndivsmiseck 397

RESULT 13
AAM09383
ID AAM09383 standard; Protein; 398 AA.
XX
AC AAM09383;
XX
DT 19-AUG-1997 (first entry)
XX
DE Human gastric lipase protein sequence.
XX
KW Duodenal; gastric; lipase; transgenic; plant; recombinant; extract;
KW food; absorption; fat; pancreatic; cystic fibrosis; exocrine; dairy;
KW hydrolysis; trans-esterification; substrate; enzyme; biofuel; human.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT Protein /note="signal peptide"
FT Protein 20..999
FT Protein /note="mature protein"
XX
PN WO9633277-A2.
XX
PD 24-OCT-1996.
XX
PF 19-APR-1996; 96MO-FR00606.
XX
PR 20-APR-1995; 95FR-0004754.
XX
PA (BIOC-) BIOCEM SA.
PA (LJOV) INST RECH JOUVEINAL.
XX
PI Baudino S, Benicourt C, Cudrey C, Gruber V, Lenee P;
PI Merot B;
XX
DR WPI; 1996-485783/48.
DR N-PSDB; AAT58916.
XX
PT Recombinant human or canine pre-duodenal lipase prodn. in transgenic
PT plants - useful for facilitating absorption of fat, as bio-catalysts
PT and for prodn. of bio-fuel
XX
PS Claim 3; Fig 5; 130pp; French.
XX
CC This is the amino acid sequence of the human pre-duodenal (i.e. gastric)
CC lipase enzyme. The sequence can be used to generate transgenic plants

```

CC Producing recombinant lipase in an enzymatically active form.
 CC Alternatively amino acids 20-23 or 20-73 (1-4 or 1-54 of the mature
 CC protein, respectively) can be deleted to form the derivatives designated
 CC delta-4 or delta-54 respectively. Plants, or their extracts, expressing
 CC the lipases or the truncated derivatives, can be used:
 CC (a) as pharmaceuticals or food to facilitate absorption of fat, either
 CC in healthy subjects or in patients with inadequate levels of fat, either
 CC gastric/pancreatic lipase e.g. those with cystic fibrosis or exocrine
 CC pancreatic insufficiency, in the elderly or in patients undergoing medical
 CC treatment which alters fat adsorption; (b) for performing industrial or
 CC agricultural reactions, e.g. in processing of fats or in the dairy
 CC industry, for hydrolysis or transesterification reactions, etc., where
 CC the plant material may provide both enzyme and substrate. The transgenic
 CC plants can also be used for biofuel production.

XX
 XX Sequence 398 AA;

Query Match 49.8%; Score 1119; DB 17; Length 398;
 Best Local Similarity 52.2%; Pred. No. 3e-105;
 Matches 210; Conservative 73; Mismatches 107; Indels 12; Gaps 2;

QY 17 MLLILVAVMQRNVNSVHMPKAV-----DPEAFMNISEIIHOGYPCXEYVATED 69
 DB 1 mwilltmasll-----svlgtlthglfgkllhpgspevtmnlsmqilywygpnueyevvled 55
 QY 70 GYLISVNRIPRGVLQPKRTGSRPVLLQHGILVGASNMISLPPNNSLGFILADAGFDVWM 129
 DB 56 gyllevrnlpygkksngntgrpvrflqhgllasaatwmslmpnsllafladagdw1 115
 QY 130 GNSRGNMRSRKHKTLSIDODEFMAFSYDEMAREFLPAVINILKGTGOKIYYGYSGT 189
 DB 116 gnsrgntwarrnllyspdsvefwaidsdemakypalidivkkgqkqihyghsgqt 175
 QY 190 TMGFIAFTMPBELAKIKMYPALPIATVKAHSGPTGKFLLPDMITGKGRKEFLYQT 249
 DB 176 tlgflafstmpslakrlktftyalapvatlvkytklmlkrlvtpslfklfgdklfyphn 235
 QY 250 RFLQOLVLYLCGOVILDOICSNIMLLGEGFTNNMNSRASVYAHTLAGTSVONILHMS 309
 DB 236 ffdgflatevcsremnlncsnalfllcgfidsknfntarldvylshnpagtsvqnmfhw 295
 QY 310 QAVNSGELRADMGSETRKNEKCNQPPVRVRYRDMVPTAMMGGODNLSPEDVYKML 369
 DB 296 qavnsqkfgaydwsypvgrmhqdsqppyyrvlamvplavwngkdlldpqpvg11 355
 QY 370 SEVNTLIYHKNIPEMAHDFWIGLDAPHRMNEIHLMOQEE 411
 DB 356 pklpnlllyhkeipfnhldfiwandapevyndlvsmseck 397

RESULT 14
 AAB66086
 ID AAB66086 standard; Protein: 398 AA.
 XX
 XX AAB66086;
 XX
 XX 30-MAR-2001 (first entry)
 XX
 DE Human lipase protein.
 XX
 XX Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
 XX central nervous system; focal brain disorder; bipolar affective disorder;
 XX global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 XX senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 XX Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 XX neuropsychiatric; psychoactive substance use; anxiety.
 OS
 XX Homo sapiens.
 XX
 XX W0200077239-A2.
 XX
 XX 21-DEC-2000.

XX
 XX 24-MAY-2000; 2000MO-US14858.
 PF
 XX
 XX 14-JUN-1999; 9905-0333159.
 PR
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 PA
 XX
 XX McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
 PI
 XX
 XX WPI; 2001-032313/04.
 DR
 XX
 XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -
 PS
 XX
 XX Disclosure; Fig 6; 359pp; English.

CC The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-FA5136 and AAF45138-FA5139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder. The present sequence is a
 CC sequence used in a sequence homology alignment with the TANGO/INTERCEPT
 CC sequences of the present invention.

XX
 XX Sequence 398 AA;

Query Match 49.8%; Score 1119; DB 22; Length 398;
 Best Local Similarity 52.2%; Pred. No. 3e-105;
 Matches 210; Conservative 73; Mismatches 107; Indels 12; Gaps 2;

QY 17 MLLILVAVMQRNVNSVHMPKAV-----DPEAFMNISEIIHOGYPCXEYVATED 69
 DB 1 mwilltmasll-----svlgtlthglfgkllhpgspevtmnlsmqilywygpnueyevvled 55
 QY 70 GYLISVNRIPRGVLQPKRTGSRPVLLQHGILVGASNMISLPPNNSLGFILADAGFDVWM 129
 DB 56 gyllevrnlpygkksngntgrpvrflqhgllasaatwmslmpnsllafladagdw1 115
 QY 130 GNSRGNMRSRKHKTLSIDODEFMAFSYDEMAREFLPAVINILKGTGOKIYYGYSGT 189
 DB 116 gnsrgntwarrnllyspdsvefwaidsdemakypalidivkkgqkqihyghsgqt 175
 QY 190 TMGFIAFTMPBELAKIKMYPALPIATVKAHSGPTGKFLLPDMITGKGRKEFLYQT 249
 DB 176 tlgflafstmpslakrlktftyalapvatlvkytklmlkrlvtpslfklfgdklfyphn 235
 QY 250 RFLQOLVLYLCGOVILDOICSNIMLLGEGFTNNMNSRASVYAHTLAGTSVONILHMS 309
 DB 236 ffdgflatevcsremnlncsnalfllcgfidsknfntarldvylshnpagtsvqnmfhw 295
 QY 310 QAVNSGELRADMGSETRKNEKCNQPPVRVRYRDMVPTAMMGGODNLSPEDVYKML 369
 DB 296 qavnsqkfgaydwsypvgrmhqdsqppyyrvlamvplavwngkdlldpqpvg11 355
 QY 370 SEVNTLIYHKNIPEMAHDFWIGLDAPHRMNEIHLMOQEE 411
 DB 356 pklpnlllyhkeipfnhldfiwandapevyndlvsmseck 397

RESULT 15
 AAB66070
 ID AAB66070 standard; Protein: 144 AA.
 XX
 XX AAB66070;
 XX
 XX 21-DEC-2000.

AC AAB6070;
 XX
 DT 30-MAR-2001 (first entry)
 XX
 DE Human TANGO 294 cytoplasmic domain.
 XX
 KW TANGO protein; INTERCEPT protein; neurological disorder;
 KW central nervous system; focal brain disorder; bipolar affective disorder;
 KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
 KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
 KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
 KW neuropsychiatric; psychoactive substance use; anxiety.
 XX
 OS Homo sapiens.
 XX
 PN MO20007239-A2.
 XX
 PD 21-DEC-2000.
 XX
 PF 24-MAY-2000; 2000MO-US14858.
 XX
 PR 14-JUN-1999; 9905-0333159.
 XX
 PA (MILL-) MILENNIUM PHARM INC.
 XX
 PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;
 DR WPI; 2001-032313/04.
 XX
 PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -
 XX
 PS Claim 8; Page 327; 359pp; English.
 XX
 CC The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.
 CC
 XX
 SQ Sequence 144 AA;
 Query Match 34.8%; Score 783; DB 22; Length 144;
 Best Local Similarity 100.0%; Pred. No. 1e-71;
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 280 MTNNMNSRSASVYAHTLAQTSVONILHMSQAVNSGELRAFDWGSSETKNLEKCNQPTPR 339
 DB 1 ntmmmsrsasyaahltlaqtsvqnllhwsqavnsgeqlrafdwgselknleknqptpr 60
 OY 340 YRYVDMVPTAMWTGGODWLSNPEDVKMLSEVTNLIYHNKIPERMAHVDPIWGLDAPHRM 399
 DB 61 YRYVDMVPTAMWTCGGQDWLSNPEDVKMLSEVTNLIYHNKIPERMAHVDPIWGLDAPHRM 120
 OY 400 YNEIHLMOQETNLSQGRCEAVL 423
 DB 121 YNEIHLMQGETNLSGRCRAVL 144

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2002, 22:11:19 ; Search time 65.9 Seconds
(without alignments)
144.444 Million cell updates/sec

Title: US-09-578-063-47

Perfect score: 2247
Sequence: 1 MLETLRQWYSHREKML.....THLMQDETNLSGRCRAVL 423

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1161	51.7	377	1 US-08-227-108-17
2	1161	51.7	377	2 US-09-073-674-17
3	1150.5	51.2	379	1 US-08-227-108-18
4	1150.5	51.2	379	2 US-09-073-674-18
5	1141.5	50.8	379	1 US-08-227-108-3
6	1141.5	50.8	379	2 US-09-073-674-3
7	1141.5	50.8	380	1 US-08-227-108-5
8	1141.5	50.8	380	2 US-09-073-674-5
9	1098.5	48.9	380	1 US-08-227-108-16
10	1098.5	48.9	380	2 US-09-073-674-16
11	105	4.7	430	1 US-08-484-105-22
12	105	4.7	430	2 US-08-484-105-22
13	102	4.5	346	1 US-08-602-359A-34
14	98.5	4.4	388	1 US-08-232-519-2
15	98.5	4.4	388	2 US-08-456-956-2
16	89.5	4.0	624	2 US-08-756-317-9
17	89	4.0	1674	2 US-08-968-542C-12
18	88.5	3.9	869	2 US-08-483-101-15
19	84.5	3.8	338	2 US-08-602-359A-40
20	84.5	3.8	935	2 US-07-707-367-2
21	84	3.7	719	4 US-09-386-607-2
22	82.5	3.7	485	2 US-08-446-803-1
23	82.5	3.7	485	2 US-08-861-837-1
24	82.5	3.7	485	2 US-08-600-908A-12
25	82.5	3.7	485	3 US-08-683-838A-12
26	82.5	3.7	485	3 US-08-600-656-1
27	82.5	3.7	485	4 US-09-170-670-1

28	82.5	3.7	485	4 US-09-170-670-7	Sequence 7, Appl1
29	82.5	3.7	485	4 US-09-193-068-1	Sequence 1, Appl1
30	82.5	3.7	485	4 US-09-193-068-7	Sequence 7, Appl1
31	82.5	3.7	485	4 US-09-183-412-1	Sequence 1, Appl1
32	82.5	3.7	485	4 US-09-183-412-7	Sequence 7, Appl1
33	82.5	3.7	485	4 US-09-264-097-7	Sequence 7, Appl1
34	82.5	3.7	485	4 US-09-354-191A-1	Sequence 1, Appl1
35	82	3.6	330	2 US-08-997-080-30	Sequence 30, Appl1
36	82	3.6	330	2 US-08-997-362-30	Sequence 30, Appl1
37	82	3.6	330	3 US-08-873-970-30	Sequence 30, Appl1
38	82	3.6	330	4 US-09-095-855-30	Sequence 30, Appl1
39	82	3.6	330	4 US-08-705-347A-30	Sequence 30, Appl1
40	82	3.6	493	1 US-07-615-448A-7	Sequence 7, Appl1
41	82	3.6	493	1 US-08-196-361-7	Sequence 7, Appl1
42	82	3.6	493	2 US-08-446-934-7	Sequence 7, Appl1
43	82	3.6	493	3 US-08-448-128-7	Sequence 7, Appl1
44	82	3.6	493	3 US-08-948-703-7	Sequence 7, Appl1
45	81.5	3.6	428	1 US-08-190-802A-29	Sequence 29, Appl1

ALIGNMENTS

RESULT 1
US-08-227-108-17
Sequence 17, Application US/08227108
Patent No. 5807726
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fauncel, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 377 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-227-108-17

Query Match 51.7%; Score 1161; DB 1; Length 377;
Best Local Similarity 57.7%; Pred. No. 2e-114;
Matches 213; Conservative 63; Mismatches 93; Indels 0; Gap:
0y 42 DPEAFNISEIIHQGCEYEYATEDGTLVSNRIPGLVOPKKGSRPVYLLQGLV 1
::: ||| ::::: |::: ||| ||||| ||||| | ||| | ||||| |||||

DB 9 NPEANNISOMITWYWGPCOEYEVTEDEGYILGVYRIPHGKNNSENIGKRPVYVLLQHLI 68
QY 102 GGASNMISLNPNNISLGIILADAGEFDYWMGNSRGNAHSRKHTLSIDODEFWASDEMAR 161
DB 69 ASATNMTANIPNNISLAFMLADAGIDVWLGNSRGNTWSRKRVYVSPDSVEFWASFDDEMAK 128
QY 162 FDLPAVINFLIQTGOEKIYVGYSGOTTGFIATFSTMPBLAOKIKMYFALADIATVKA 221
DB 129 YDLPATINFLVQGTGEEKIHVGHSGOTTIGFIATFSTNPLAKIKITFYALAVATVKKY 168
QY 222 KSPGTFFLLPDMKIGLBEKKEFLYOTRPLROLVYILCGOVLIDICSNIMLLGFTNT 261
DB 189 QSPBKRSIFPTFLFKLMFKMKMFLPHTYFDDFLGTEVCSEVLDLCSNTLPIFCGFDR 248
QY 282 NNNMSRASYAAHTLAGTSVONILHMSQAVNSGELRAFDMGSETKNLEKCNOPTRVRY 341
DB 249 KNLVSRFDVYLGHNPAGTSVODFLHMAQLVRSKGFQAFWGSPOMLHYNOKTPREYD 308
QY 342 VRDVTPTAMTGGODMLSNPDEVKMLSEVTMLIYHKNIPEMAHYDFIWLDAPIHRMYN 401
DB 309 VSAMTVPVAVWNGNDILADPDQVAMLLPKLSNLPFKHELAVNHDFIWMADAPQEVYN 368
QY 402 EIIHMQOE 410
DB 369 EMISMAED 377

RESULT 2
US-09-073-674-17
Sequence 17, Application US/09073674
Patent No. 5998189
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Warner-Lambert Company
STREET: 2800 Plymouth Road
CITY: Ann Arbor
STATE: Michigan
COUNTRY: U.S.A.
ZIP: 48105

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,674
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Crissey, Todd M.
REGISTRATION NUMBER: 37,807
REFERENCE/DOCKET NUMBER: 5072-DI-66-TWC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 734 622-7530
TELEFAX: 734 622-1553
TELEX:
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 377 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-073-674-17

Query Match 51.7%; Score 1161; DB 2; Length 377;
Best Local Similarity 57.7%; Pred. No. 2e-114;

Matches 213; Conservative 63; Mismatches 93; Indels 0; Gaps 0;
QY 42 DPEAFNHSIFLHOHGPCEYEVAATEDGYILSVNRPBGLVQPKTGSPVYVLLQHLI 101
DB 9 NPEANNISOMITWYWGPCOEYEVTEDEGYILGVYRIPHGKNNSENIGKRPVYVLLQHLI 68
QY 102 GGASNMISLNPNNISLGIILADAGEFDYWMGNSRGNAHSRKHTLSIDODEFWASDEMAR 161
DB 69 ASATNMTANIPNNISLAFMLADAGIDVWLGNSRGNTWSRKRVYVSPDSVEFWASFDDEMAK 128
QY 162 FDLPAVINFLIQTGOEKIYVGYSGOTTGFIATFSTMPBLAOKIKMYFALADIATVKA 221
DB 129 YDLPATINFLVQGTGEEKIHVGHSGOTTIGFIATFSTNPLAKIKITFYALAVATVKKY 168
QY 222 KSPGTFFLLPDMKIGLBEKKEFLYOTRPLROLVYILCGOVLIDICSNIMLLGFTNT 261
DB 189 QSPBKRSIFPTFLFKLMFKMKMFLPHTYFDDFLGTEVCSEVLDLCSNTLPIFCGFDR 248
QY 282 NNNMSRASYAAHTLAGTSVONILHMSQAVNSGELRAFDMGSETKNLEKCNOPTRVRY 341
DB 249 KNLVSRFDVYLGHNPAGTSVODFLHMAQLVRSKGFQAFWGSPOMLHYNOKTPREYD 308
QY 342 VRDVTPTAMTGGODMLSNPDEVKMLSEVTMLIYHKNIPEMAHYDFIWLDAPIHRMYN 401
DB 309 VSAMTVPVAVWNGNDILADPDQVAMLLPKLSNLPFKHELAVNHDFIWMADAPQEVYN 368
QY 402 EIIHMQOE 410
DB 369 EMISMAED 377

RESULT 3
US-08-227-108-18
Sequence 18, Application US/08227108
Patent No. 5807726
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-227-108-18


```

; INFORMATION FOR SEQ ID NO: 3
;
; SEQUENCE CHARACTERISTICS:
;
;   LENGTH: 379 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
US-08-227-108-3

```

Query Match	50.8%	Score 1141.5	DB 1	Length 379
Best Local Similarity	56.5%	Pred. No. 2.4e-112		
Matches 210	Conservative 67	Mismatches 90	Indels 5	Gaps 3

Qy	37	PTVDEAFAMNISEIIIOHOGYPOEEVEATEOGIIVSIRIRGLVQPKKTSRPVLL	96
Db	7	PT---NEVTMNISOMTITGTYPAEEVEVTEDEYIIIGDIRIPGRKNSNIGRRPAFL	63
Qy	97	QHGLVGASNMISLNPNSLGFILADAGFDVMGNSRGNAHSRKHKLSIDODEFNAFSY	156
Db	64	QHGLASATIMWISLPPNLSLAFILADAGDYVILGNSGNTMARRNLIYSPDSVEFNAFSF	123
Qy	157	DEMARFDLPVAVINFILOKTGOEKIYYUYGSGTTMGFIASFYMEBELAQKTKMTFALAPIA	216
Db	124	DEMAKYLPLPATIDFILLKTGODKLHHYGHSGQTTIGFIASFSTPKLAKRIKTFYALAPVA	183
Qy	217	TYKAKSPGRKFLLPMMIKGLGKKKEFLYQTFELNO-LVYLICGVILDOIICSNMML	275
Db	184	TYKTEFLTKMLKMLPVSFELKLLFGNKIF-YPHHFFDQFLATEVCSREYDILCSNALFTI	242
Qy	276	LGCFENTNMNMSRASYAAHTLAGTSYONLILHMSOAVNSGELRAFDGSEFTKMLEKNOP	335
Db	243	ICGDFIDMLNMSRLDYVLSHNPAGTSYQVNLILHMSQAKSKFQAFDMGSPSYQNMNHHOS	302
Qy	336	TPARYRVRDMTVPYAMTGGODWLMSNEDYVKMLISEVTNLITYHKNIPEMAHVDFTWGLDA	395
Db	303	MPRYNLTDHMPYPAVYVNGDNLADBDVLLSLKPLNLIYHRKIPRYNHDFIWAMD	362
Qy	396	PHKMYNELIHLM 407	
Db	363	PQAVYNELVISM 374	

RESULT 6
 US-09-073-674-3
 : Sequence 3, Application US/09073674
 : Patent No. 596189
 :
 : GENERAL INFORMATION:
 : APPLICANT: Blanchard, Claire
 : APPLICANT: Benicourt, Claude
 : APPLICANT: Juntten, Jean-Louis
 : TITLE OF INVENTION: Recombinant Dog Gastric Lipase
 : NUMBER OF SEQUENCES: 21
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Warner-Lambert Company
 : STREET: 2800 Plymouth Road
 : CITY: Ann Arbor
 : STATE: Michigan
 : COUNTRY: U.S.A.
 : ZIP: 48105
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/073,674
 : FILING DATE:
 : CLASSIFICATION:
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Cilssey, Todd M.
 : REGISTRATION NUMBER: 37,807
 : REFERENCE/DOCKET NUMBER: 5072-D1-66-TWC
 : TELECOMMUNICATION INFORMATION:

TELEPHONE: 734 622-7530
TELEFAX: 734 622-1553
TELEX:
INFORMATION FOR SEQ. ID NO.:
SEQUENCE CHARACTERISTICS:
LENGTH: 379 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-073-674-3

Query Match	50.8%	Score 1141.5;	DB 2;	Length 379;
Best Local Similarity	56.5%	Pred. No. 2.4e-112;		
Matches 210; Conservative	67;	Mismatches 90;	Indels 5;	Gaps 3;

QY	37	PTKRVDEAFMNISSEITIOHOGYCEEEYAVTEDEGYILSVNRIPRGVLQVPPKTSRPEVLL	96
Db	7	PT---NEVTMNISSQMTTMYGVPRAEEVEVUTEDGYILGIDRIYGRKNSNINIRRPVAFLL	63
QY	97	OHGLVAGSASWISNLPPNNSLGFLLADAGFDVMMGNSRKNMRSKHKLTSLIDODEFMAFSY	156
Db	64	OHGLLASATWISNLPNNSLAFLLADAGDYWVLCNSRGNTRWARRNLYSPDSVEFMAFSF	123
QY	157	DENAERFLPRAVINFILOKTGOEKIYYUYGYOQGTMMGTIASTYMBELAQKKMFALAPTA	216
Db	124	DEMAKYLPLPATIDFLIKKTQODKILHYHGSQGTIGFIAESTPDKLARKIKETVYALAPVA	183
QY	217	TVTKAKSPGRFLPLPMMYIKGLEGGKEFYQTFRFLNQ-LVYILCGOVLIDQOCSNIMLL	275
Db	184	TVKTKTEELLKMLKMYBSEFLTKLIFGNKIF-YPHHFFQFLATEYCSRETYDLLCSNRLFI	242
QY	276	LGEFNTNNMMSRASVYAAHTLAGTSYQNTILHMSQAVNSGELRAFDMGSETSKMLEKONP	335
Db	243	ICGEDTNNLMNSRLDYVLSHNPAGTSYQNTILHMSQAVKSKFQAFDMGSPVQMMMYHQS	302
QY	336	TPVRIYVRDNTVPTAMTGGODMLSNPDEYKMLISEVTNLITKHNIPPEMAHNVDFMGDLA	395
Db	303	MPEYNTLDMHVPYAVVWNGNDLADPHVDLLSKLPIYHRRKIPRYNHDFIWAMDA	362
QY	396	PHRYNTEILHLM 407	
Db	363	POAYINELVSM 374	

RESULT 7
US-08-227-108-5
Sequence 5, Application US/08227108
Patent No. 5807726
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Julien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipa
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fannucci, Allan A.

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-Apr-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-227-108-16

Query Match 48.9%; Score 1098.5; DB 1; Length 380;
Best Local Similarity 54.6%; Pred. No. 8.4e-108;
Matches 202; Conservative 69; Mismatches 98; Indels 1; Gaps 1;

QY 43 PEAFNMISEIIOHOGPCEEYEATEDGTYLSVNRIRGLVOPKKTGRPVVLLQHGVLG 102
10 PEVTMNSQMITYWGYNENEYETEDGYLLEVNRIPIYGRKNSGNTQGRPVVFLQHGLLA 69
DB 10 PEVTMNSQMITYWGYNENEYETEDGYLLEVNRIPIYGRKNSGNTQGRPVVFLQHGLLA 69
QY 103 GASWISNLNNSISGLFLAAGFDVWNGSRGNAMSKRHKTLSDIODEFV-ARSYDEMAR 161
70 SATNMISNLNNSISGLFLAAGFDVWNGSRGNAMSKRHKTLSDIODEFV-ARSYDEMAR 129
DB 103 GASWISNLNNSISGLFLAAGFDVWNGSRGNAMSKRHKTLSDIODEFV-ARSYDEMAR 161
QY 162 FDPYAVINFILOKTGOEKIYVGSOGTGMGFIASFSTMPBLAOKIKYFALAPATYKHA 221
130 YDLPATIDFIVKKTGOQLHYVHSGSTTIGFIASFSTMPBLAOKIKYFALAPATYKHT 189
DB 130 YDLPATIDFIVKKTGOQLHYVHSGSTTIGFIASFSTMPBLAOKIKYFALAPATYKHT 189
QY 222 KSPGTEKLLPDMMIKGLFGKKEFLYOTRFLRQLYVILCGOVIIDQCSNIMLLGCFNT 281
190 KSLINKLRFVPOSLEFKFIYFNHFDQFLATEVCSREMLNLLCSNALFLICGFDS 249
DB 190 KSLINKLRFVPOSLEFKFIYFNHFDQFLATEVCSREMLNLLCSNALFLICGFDS 249
QY 282 NNNMNSRASYAAHTLAGTSVONLHWSQAVNSGELAFPMGSEPTKNLEKNOPTPVRYR 341
250 KNFTSRDLVYLHSHNPAGTSVONHFWTOAVKSGKFOAYDMGSPVONRHMHDOSQPPYYN 309
DB 250 KNFTSRDLVYLHSHNPAGTSVONHFWTOAVKSGKFOAYDMGSPVONRHMHDOSQPPYYN 309
QY 342 VRDQTVTAMTGGODWLSNPEDVKMLLSEVTNLIYHKNIPEMAHVDPIWGLDAPHRMYN 401
310 VTAMNVPYIAVWNGSKDLADPDVGLLPRLPLNLIYHKEIPFYVHDLFIAMADAPQEVYN 369
DB 310 VTAMNVPYIAVWNGSKDLADPDVGLLPRLPLNLIYHKEIPFYVHDLFIAMADAPQEVYN 369
QY 402 EIHLMOOE 411
370 DIVSMISEDK 379
DB 370 DIVSMISEDK 379

RESULT 10
US-09-073-674-16
Sequence 16, Application US/09073674
Patent No. 5998189
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Julien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Warner Lambert Company
STREET: 2800 Plymouth Road
CITY: Ann Arbor
STATE: Michigan
COUNTRY: U.S.A.
ZIP: 48105
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,674
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Crissey, Todd M.
REGISTRATION NUMBER: 37,807
REFERENCE/DOCKET NUMBER: 5072-DI-66-TMC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 734 622-7530
TELEFAX: 734 622-1553
TELEX:
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-073-674-16

Query Match 48.9%; Score 1098.5; DB 2; Length 380;
Best Local Similarity 54.6%; Pred. No. 8.4e-108;
Matches 202; Conservative 69; Mismatches 98; Indels 1; Gaps 1;

QY 43 PEAFNMISEIIOHOGPCEEYEATEDGTYLSVNRIRGLVOPKKTGRPVVLLQHGVLG 102
10 PEVTMNSQMITYWGYNENEYETEDGYLLEVNRIPIYGRKNSGNTQGRPVVFLQHGLLA 69
DB 10 PEVTMNSQMITYWGYNENEYETEDGYLLEVNRIPIYGRKNSGNTQGRPVVFLQHGLLA 69
QY 103 GASWISNLNNSISGLFLAAGFDVWNGSRGNAMSKRHKTLSDIODEFV-ARSYDEMAR 161
70 SATNMISNLNNSISGLFLAAGFDVWNGSRGNAMSKRHKTLSDIODEFV-ARSYDEMAR 129
DB 103 GASWISNLNNSISGLFLAAGFDVWNGSRGNAMSKRHKTLSDIODEFV-ARSYDEMAR 161
QY 162 FDPYAVINFILOKTGOEKIYVGSOGTGMGFIASFSTMPBLAOKIKYFALAPATYKHA 221
130 YDLPATIDFIVKKTGOQLHYVHSGSTTIGFIASFSTMPBLAOKIKYFALAPATYKHT 189
DB 130 YDLPATIDFIVKKTGOQLHYVHSGSTTIGFIASFSTMPBLAOKIKYFALAPATYKHT 189
QY 222 KSPGTEKLLPDMMIKGLFGKKEFLYOTRFLRQLYVILCGOVIIDQCSNIMLLGCFNT 281
190 KSLINKLRFVPOSLEFKFIYFNHFDQFLATEVCSREMLNLLCSNALFLICGFDS 249
DB 190 KSLINKLRFVPOSLEFKFIYFNHFDQFLATEVCSREMLNLLCSNALFLICGFDS 249
QY 282 NNNMNSRASYAAHTLAGTSVONLHWSQAVNSGELAFPMGSEPTKNLEKNOPTPVRYR 341
250 KNFTSRDLVYLHSHNPAGTSVONHFWTOAVKSGKFOAYDMGSPVONRHMHDOSQPPYYN 309
DB 250 KNFTSRDLVYLHSHNPAGTSVONHFWTOAVKSGKFOAYDMGSPVONRHMHDOSQPPYYN 309
QY 342 VRDQTVTAMTGGODWLSNPEDVKMLLSEVTNLIYHKNIPEMAHVDPIWGLDAPHRMYN 401
310 VTAMNVPYIAVWNGSKDLADPDVGLLPRLPLNLIYHKEIPFYVHDLFIAMADAPQEVYN 369
DB 310 VTAMNVPYIAVWNGSKDLADPDVGLLPRLPLNLIYHKEIPFYVHDLFIAMADAPQEVYN 369
QY 402 EIHLMOOE 411
370 DIVSMISEDK 379
DB 370 DIVSMISEDK 379

RESULT 11
US-08-484-105-22
Sequence 22, Application US/08484105
Patent No. 5589341
GENERAL INFORMATION:
APPLICANT: STILLMAN, Bruce
APPLICANT: BELL, Stephen P.
APPLICANT: KOBAYASHI, Ryuji
APPLICANT: RINE, Jasper
APPLICANT: FOSS, Margit
APPLICANT: MCNALLY, Patricia J.
APPLICANT: LAURENSEN, Patricia
APPLICANT: HERSKOWITZ, Ira
APPLICANT: LI, Joachim J.

APPLICANT: GAVIN, Kimberly
TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,105
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-105-22

Query Match 4.7%; Score 105; DB 1; Length 430;
Best Local Similarity 21.9%; Pred. No. 0.013;
Matches 61; Conservative 42; Mismatches 87; Indels 88; Gaps 13;
QY 85 PKKTGRPVYLLQHLGVGASNMISNLPNN-----SLGF-----ILADAGFDVW--MGN 131
DB 25 PEKGRSR-----QKTKGRKNASRNLSNLEEDLQGFEDETYSMAQSALENFMQGX 78
QY 132 S-----RGNAMSRKH-----TLSDODEFMAFSYDENARFDLPVINFILQKTOEKTY 181
DB 79 SASERMMNAKSRGRAGNGNTEIEEDDEISNATDFTKCDLPGLRNYITRKDKNTEFEK 138
QY 182 YVGSQGTMGFIASFTEPDLAQKIKMFALAPITVYKHAASPGTKFLLPDMIKGLFG 241
DB 139 RLEHLADNDFG-----KWKLYLAG-----FNILLHGVGS 168
QY 242 KKEFLYOTRFLQOLVILYCGOYILDOICSNIMLLGFGNTN---NMNRSASVYAHTLA 298
DB 169 KRDLV--TEFENELSDYTYMRVDARKDGLNVKVLGAINEMKLNCKNVRG----- 217
QY 299 GTSVQNILHWSQAV---NSGEL-----RAFDWGESE 325
DB 218 ---QSTISWARSIRKKNMSOOLLILIIDNIEAPDWRSD 251

RESULT 12
US-08-484-106-22
Sequence 22, Application US/09484106
Patent No. 5614518
GENERAL INFORMATION:
APPLICANT: STILLMAN, Bruce
APPLICANT: BLET, Stephen P.
APPLICANT: KOBAYASHI, Ryuji
APPLICANT: RINE, Jasper
APPLICANT: JOSS, Markit
APPLICANT: GUNALY, Francis J
APPLICANT: LAURINSON, Patricia

APPLICANT: HERSKOWITZ, Ira
APPLICANT: LI, Joachim J
APPLICANT: GAVIN, Kimberly
TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,106
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-106-22

Query Match 4.7%; Score 105; DB 1; Length 430;
Best Local Similarity 21.9%; Pred. No. 0.013;
Matches 61; Conservative 42; Mismatches 87; Indels 88; Gaps 13;
QY 85 PKKTGRPVYLLQHLGVGASNMISNLPNN-----SLGF-----ILADAGFDVW--MGN 131
DB 25 PEKGRSR-----QKTKGRKNASRNLSNLEEDLQGFEDETYSMAQSALENFMQGX 78
QY 132 S-----RGNAMSRKH-----TLSDODEFMAFSYDENARFDLPVINFILQKTOEKTY 181
DB 79 SASERMMNAKSRGRAGNGNTEIEEDDEISNATDFTKCDLPGLRNYITRKDKNTEFEK 138
QY 182 YVGSQGTMGFIASFTEPDLAQKIKMFALAPITVYKHAASPGTKFLLPDMIKGLFG 241
DB 139 RLEHLADNDFG-----KWKLYLAG-----FNILLHGVGS 168
QY 242 KKEFLYOTRFLQOLVILYCGOYILDOICSNIMLLGFGNTN---NMNRSASVYAHTLA 298
DB 169 KRDLV--TEFENELSDYTYMRVDARKDGLNVKVLGAINEMKLNCKNVRG----- 217
QY 299 GTSVQNILHWSQAV---NSGEL-----RAFDWGESE 325
DB 218 ---QSTISWARSIRKKNMSOOLLILIIDNIEAPDWRSD 251

RESULT 13
US-08-602-359A-34
Sequence 34, Application US/08602359A
Patent No. 5942430
GENERAL INFORMATION:
APPLICANT: ROBERTSON, Daniel E.
APPLICANT: MURPHY, Dennis
APPLICANT: REID, John
APPLICANT: MAFIA, Anthony
APPLICANT: LINK, Steven

```

      ZIP: 10020
      COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.25
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/232,519
      FILING DATE:
      CLASSIFICATION: 435
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: JP 5-96286
      FILING DATE: 22-Apr-1993
      ATTORNEY/AGENT INFORMATION:
      NAME: Haley Jr, James F
      REGISTRATION NUMBER: 27,794
      REFERENCE/DOCKET NUMBER: SIGN-4
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (212) 596-9000
      TELEFAX: (212) 596-9090
      TELEX: 14-8367
      INFORMATION FOR SEQ ID NO: 2:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 388 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      US-08-232-519-2

```

	Query Match	4.4%	Score 98.5;	DB 1:	Length 388;
	Best Local Similarity	20.2%	Pred. No. 0.052;		
	Matches	81;	Conservative	64;	Mismatches 142; Indels 113; Gaps 20;
OY	38	TKAVDPEAFNNISLIIHQGIPCEEYE--VATEDGIILSVNRIRPG--LVOPKKTGSR--	91		
Bd	84	TPVDIEFI-----KNFHAENPDMAATEDASAGEPKVTGTGEHTLNVDGTQEG	135		
OY	92	-PVVLHGHLVGAGSANWISLPNNSLCFLILDAGEDVMNGSGNAWSRKH--KTLSD	147		
Bd	136	TRIVLY-HGFSGDISNMU--LTODALAERRVIAPDL----PGHAGSSKNVGTGLAF-	186		
OY	148	QDEFWAPSIDEMAFDDLPAVINFLIQTCKQEKIYYVGSOGCTTWFIAFTSPBELLAKIK	207		
Bd	187	-----LAGVSELSLOTKIEAHAYVHSLG---GGIALTLRLDDPOVA	227		
OY	208	MYFLAPIATYVKAHAKSPQTKFLLPDMMIKGLFGKEKELYTQTRLROLVITY--LCGOVI	264		
Bd	228	SLNLAPAGL--KGDVADFI-----SAFDSESSRDRAVLQMLVYRKALVGRKM	276		
OY	265	LDOCSNIMLLGGFNTNNMMSRASVYAHTTLAGTSTQNILHMGAOAVNSGLRAFMDGS	324		
Bd	277	VDAV-----LRARILLGAR--ALHYIAACPNN-----GHQADDLSVLAGA	317		
OY	325	ETKLKCNQPTPYRYRVKMDTVPTAMTQGSDMLSNEDYKMLLSEVTNLTYHK-NIP	382		
Bd	318	ET-----PTQIWGEKDELTSVNAAGLDPIVPVFYEETGHLP	356		
OY	383	EMAHVDFWIGLDAPHMYNEIIHL-MQOEENLSGRCSEA	421		
Bd	357	OLEHATDV-----NKAILALEFKDPPEALSHARDWA	386		
RESULT	15				
US-08-456-956-2					
: Sequence 2, Application us/08456956					
: Patient No. 5686585					
GENERAL INFORMATION:					
APPLICANT: Kageyama, Bunji					
APPLICANT: Nakae, Masanori					
APPLICANT: Yagii, Shigeo					
TITLE OF INVENTION: No. 5686285Dornane Type Ester Hydrolase					
NUMBER OF SEQUENCES: 2					

Search completed: January 30, 2002, 23:01:20
Job time: 3001 sec

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,956
FILING DATE: 01-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-96286
FILING DATE: 22-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/232,519
FILING DATE: 22-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr, James F
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: SHGN-4 DIV 1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)596-9000
TELEFAX: (212)596-9090
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 388 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-456-956-2

Query Match      4.4%; Score 98.5; DB 1; Length 388;
Best Local Similarity 20.2%; Pred. No. 0.052;
Matches 81; Conservative 64; Mismatches 142; Indels 113; Gaps 20;

QY 38 TKAVPEAFNINISIIHQGPCEYE--VAEDGYILSVNRIPRG--LVQPKKTGSR-- 91
   ||| ||| : : : ||| : : : : : : : : : : : : : : : : : :
Db 84 TPVDIDIAFI-----KNFHAENPODAATEASAGEPKQVTGHEHTLVNRDVGTOEG 135

QY 92 -PVYLLQGLVGSASNTSNPNNSLGFILADAGFDVYMGNSRGNAMSRKH--KTLSTD 147
   ||| ||| ||| ||| : : : ||| : : : : : : : : : : : : : : : : : :
Db 136 TPVLYV-HGFGDISNML--LTODALAERRVIAFDL-----PGHGASSKNVGTGLAF- 186

QY 148 QDEFWAFSYDEMARFDELPAVINFILOKTGOKIYVYGSGQTMGFIASFMPLEAQKIK 207
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 -----LAGVSELLQTLKIEKAHVGHSLG---GGIALTLIRDPDOVA 227

QY 208 MYFALAPIATVKAHAKSPGKFLLPDMMIKGLFGKKEFLYQTRFLRLQVLY--LCGQVI 264
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 228 SLNLLAPAGL---GKDVNADFI-----SAFVDESSRDMKAVLQMLVYKKAIVGRKM 276

QY 265 LDQICSNMMLLGGFNTNMMMSRASVYAATLACTSVQNLHMSQAVNSGELRAFDWGS 324
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 277 VDAV-----LRARRLDGARD--ALHVIACFPN-----GHQADDLHSLVLAGA 317

QY 325 ETKNLKCNQPTPVRYRVRDVTPTAMWTGQDMLSNPEDEVKMLLSEVTNLIYHK--NIP 382
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 ET-----PQITWCKEDELISVSNAGLPDVIPTVYEETGHP 356

QY 383 EWAHVDFTWGLDAPHRMYNEIHL-MQGEETNLSGGRCEA 421
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 357 QLEHATDV-----NKAIALFVKDPEALSMARMDA 386
```

[illegible]

```

RESULT      2
G01416
lysosomal acid lipase - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 18-Jun-1999
C:Accession: G01416
R:Du, H.
submitted to the EMBL Data Library, April 1994
A:Reference number: G06919
A:Accession: G01416
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-339 <DUX>
C:Cross-references: EMBL:008464; NID:g505052; PIDN:AAB0328.1; PID:g505053
C:Superfamily: triacylglycerol lipase, lingual
C:Keywords: glycoprotein

```

Query Match	57.4%	Score 1290	DB 2	Length 399
Best Local Similarity	61.0%	Pred. No. 6.1e106		
Matches 242	Conservative 59	Mismatches 94	Indels 2	Gaps 1

QY	15	MEMLLILVAVMEFGNANVSHMPTK	ADVPEAFMINISEIIOHGVCCEFEVATDEGCI	72
Db	1	MKMFLELVYCVLVTMLHSEGSCKKL	ADVPETNMNVSEITISWGFSEEXLVETEDGYI	60
QY	73	LSVNRIPIRGVLPCKTGSRPVLLH	LOGVAGASNMISNLPNNSIGFILLADAGEVMMGNS	132
Db	61	LCNRIPIHGRKNHSDKPKRPVPLQ	HOHLADSSNMVYTNLANSISIGFILLADAGEVMMGNS	120
QY	133	RGNAHSKKHKTLSIDODEFAFST	DEMAFIDLPAVINFLQKTGOEKIYVYGVSQGTJMG	192
Db	121	RGNTWSKKHKTLSVODEFAFST	DEMAKXDLPASINFILNTKGQEQVYVYVHGSQGTIG	180
QY	193	FIAFSTPELAKRKMEFALPATVY	HAHSPGTFLLPDDMIKGLFGKKEFLYQTRFL	252
Db	181	PIAFSQPELAKRKMEFALGPV	ASAVFCTSPMAKIGLRIDHLIKDFEGDEFLPOSATL	240
QY	253	ROLVILYLCGOVILIDOLCSN	MLLGGFNTNMNMSRASVYAAHTLAGTSVQNTLHMSQAV	312
Db	241	KWLGTHCHTVLILKELOGN	ICFLCGENENLMSKSRVDYTTISPACTSVQNTLHMSQAV	300
QY	313	NSGELRAFDMGSETKNLKCN	QPTPVARYRVDMFTVPTAMTNGQDMLSNEDVKMLSEV	372
Db	301	KFOKFOAFDMGSSAKNFHY	NOSYPRPYNKDMLVPTAVYMSGDHMLADVYDNIILTOI	360
QY	373	TNLIYHKNIPEMAHVDPI	MGDLAPHRKYNNIILMOQ	409
Db	361	TNLIYHESIPMEHDLPI	WGLDAPWRKYNNIILMR	397

RESULT 3
LIRTT
triacylglycerol lipase (EC 3.1.1.3) precursor, lingual - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 11-Jun-1999
C:Accession: A23045
R:Docterly, A.J.P.; Bodmer, M.W.; Angal, S.; Vergar, R.; Rylare, C.; Lowe, P.A.; L.
Nucleic Acids Res. 13, 1891-1903, 1985
A:Title: Molecular cloning and nucleotide sequence of rat lingual lipase cDNA.
A:Reference number: A23045; MUID:85215587
A:Accession: A23045
A:Molecule type: mRNA
A:Residues: 1-395 <DOC>
A:Cross-references: GB:X02309; NID:g56595; PIDN:CAA26179.1; PID:g56596
A:Experimental source: strain Sprague-Dawley
A>Note: The partial sequence of the mature protein from a different, unspecified st
C:Comment: This acid-stable lipase is secreted by the serous (von Ebner's) glands at
C:Superfamily: triacylglycerol lipase, lingual
C:Keywords: carboxylic ester hydrolase; glycoprotein; lipid digestion; saliva; serou
F:1-18/Domains: signal sequence status predicted <Sig>
F:19-395/Product: triacylglycerol lipase, lingual #status predicted <MP>
F:33, 68, 98, 184, 270/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match	51.8%	Score 1164.5	DB 1	Length 395
Best Local Similarity	55.9%	Pred. No. 7.1e-95		
Matches 217	Conservative 69	Mismatches 108	Indels 1	Gaps 1

QY	17	MLLLILVAYMqQ -NNVNSVNHPTKAVDPDEAFNMSEIIHOHQYCCXEYVATEBGYTLV	75
Dp	1	MLLLILVISTJTEGGAGHLEKGLCKLPNDPEANMNSOMITWYGPQCEYEYVTEBGYTLV	60
QY	76	NRIBRGIVQPKKTSRPPVLLLOHGLVYGASNMWISNLSGFIILADAGFDVMMGNSRGN	135
Dp	61	YRIFHGKNNNSNIKRPPVYVLOHLLISATYMIANLNPNSLAFPLADAGDVMGNSRGN	120
QY	136	AMSRKHKTLSDIDQEFMAFSTEDARFDLPVYINFIQKTGOEKIYYVYSOGTWGFI	195
Dp	121	TWMSRKNVYSPDSVEPFAFSEDEAKKADLPATINFIYQKTGOEKIHVYHSGOCTTIGFI	180
QY	196	FSYMPLEAOKIKMYEFALAPATVYTHAKSPGKFKLLDDMMIKIGLFGKKEFLYOTRFLROL	255
Dp	181	FSYMPLEAKKIKTYALAPAVYVYVYTOPSLKISIFTEFLKMLFGKMKMLPRTYFDLFL	240
QY	256	VYVLCGOVYIQQIOLISNIMLLGSGFNTNMNMSRASVYAATFLAGTSVONILHNSOAVNSG	315
Dp	241	GTEVCSREYVDDLCSNMLTFEGSGDKKNLVNSRDVYLGNHPAGTSVQDFLHMAQLYVRSG	300
QY	316	ELRAFDWGSSETKNLKCNQPTPVRYRVYRDMKTVPYAMTSGQDWLSNEDVYKMLLSEVTNL	375
Dp	301	KFOAFNNGSPSONMLHYNOKTRPEYDASVATVPYAVVNGNDLILADPODVAMLLPKLSNL	360
QY	376	IYHKNIPEMAHVDFIWMELDAPHRKYNNIILHMOE	410
Dp	361	LEHKEILAYNHLDIYAMADAPQEVYENIIMIAEMD	395

RESULT 4

JC4017

tricylglycerol lipase (EC 3.1.1.3) PGE precursor - bovine

N:Alternate names: pregastric esterase

C:Species: Bos primigenius taurus (cattle)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: JC4017, S64678

R:Timmermans, M.Y.U.; Teuchy, H.; Kupers, L.P.M.

A:Title: The cDNA sequence encoding bovine pregastric esterase.

A:Reference number: JC4017, MUID:95011625

A:Accession: JC4017

A:Molecule type: mRNA

A:Residues: 1-397 <TIM>

A:Cross-references: GB:126319, NID:g600756, PID:g600757

A:Experimental source: tongue

R:Timmermans, M.Y.J.; Reekmans, G.; Teuchly, H.J.H.; Kupers, L.P.M.

Biochem. J. 314, 931-936, 1996

A:Title: Inhibition studies on calf pregastric esterase: the enzyme has no functional th

A:Reference number: 564678; MUID:96177869

A:Accession: 564678

A:Molecule type: protein

A:Residues: 24-40;248-253 <TIM>

C:Comment: Pregastric esterase is a major fat-digesting enzyme.

C:Genetics:

A:Gene: pge

C:Superfamily: triacylglycerol lipase, lingual

C:Keywords: blocked amino end; carboxylic ester hydrolase; glycoprotein

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-397/Product: pregastric esterase #status predicted <MAT>

F:33,270,326/Binding site: carbohydrate (asn) (covalent) #status predicted

F:11/Active site: Ser #status predicted

Query Match 50.7%; Score 1139; DB 1; Length 397;

Best Local Similarity 53.7%; Pred. No. 1,3e-92;

Matches 212; Conservative 68; Mismatches 113; Indels 2; Gaps 2;

QY 18 WLLIIVAVMFORNVSVHMPTRKAVDPEAFNISEIIIOHGYPCOEYEVATEDGYILSVNR 77

Db 3 WLLVTVCFIHMGNAFCFLGKIAKNPEASMNVSQMSYMGYSEMKKVTIADGYILQVYR 62

QY 78 IPRGLVOPKRTGSRPVVLLQHLVGASNMISNLPNNSIGFTIADAGFDVWNGNSGNM 137

Db 63 IPHGKNNMHLQORPVVFLQHLGASATNMISNLPNNSIGFTIADAGFDVWNGNSGNM 122

QY 138 SKKHTSLSDQEFMAFSYDEMARFDLPVAVINFILOKGOEKIYVYGSGGTMGFIAS 197

Db 123 AOEHLVSPDSPEFMAFSYDEMAEYDLPSTIDFIRTOQKLUHYHSGGTGFIAS 182

QY 198 TPELAIKIKMFALAPITVYKHAASPGKFFLLPDMIKIGFGKKEFLYQEFRLQ-LV 256

Db 183 TSPFLAEKIKVYALAPVATVYKTSLEFKLALIRHFLKLIIFGDMF-YPHFLEQLG 241

QY 257 IYLCQVILDOICSNIMLLGFPNTNNMNSRASVYAAHTLAGTSYONILHMSQAVNSGE 316

Db 242 VEMCSEETDVLCKNALFATITGVNKNFMNSRLDYIAHPNAGTSYONILHMSQAVNSGE 301

QY 317 LRAFMGSEFTKMLEKNOPTPVRYRYRDMTPTAMTGGQDMLSPNEDVKMLSEVTNLI 376

Db 302 FOAFPMGAYQNLHMYHOFPTPIYTLTANVPPIAVMSADNDLADPODVLKSLNLI 361

QY 377 YHKNIPEMAHVDPIWGLDAPHRMYNEIILHMOOE 411

Db 363 YHKEIPNVNHLDFIWMADAPQEVYNEIYSLMAEDK 396

RESULT 5

S07145

triacylglycerol lipase (EC 3.1.1.3) precursor, gastric - human

C:Species: Homo sapiens (man)

C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 18-Jun-1999

C:Accession: S07145; S27102; S04942

R:Boemer, M.W.; Angel, S.; Yarrinton, G.T.; Harris, T.J.R.; Lyons, A.; King, D.J.; Piero

Biochim. Biophys. Acta 909, 237-244, 1987

A:Title: Molecular cloning of a human gastric lipase and expression of the enzyme in yea

A:Reference number: S07145; MUID:87299724

A:Accession: S07145

A:Molecule type: mRNA

A:Residues: 1-398 <BOD1>

A:Cross-references: EMBL:X05997; NID:g31771; PIDN:CAA29413.1; PID:g758063

A:Accession: S27102

A:Molecule type: protein

A:Residues: 20-24,'X',26,'X',28-33,'X',35-42,'X',44,'X',46-48,'X',50,'X',52,'X',54-56 <E

A:Note: It is uncertain whether Met-1 or Met-7 is the initiator

R:Bernbeck, S.; Blaueberg, L.

Eur. J. Biochem. 182, 495-499, 1989

A:Title: Human gastric lipase. The N-terminal tetrapeptide is essential for lipid bindin

A:Reference number: S04942; MUID:89325292

A:Accession: S04942

A:Molecule type: protein

A:Residues: 20-45 <BER>

C:Superfamily: triacylglycerol lipase, lingual

C:Keywords: carboxylic ester hydrolase; extracellular protein; glycoprotein; lipid di

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-398/Product: triacylglycerol lipase, gastric #status experimental <MAT>

F:34,99,185,271/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 49.8%; Score 1119; DB 2; Length 398;

Best Local Similarity 52.2%; Pred. No. 7.5e-91;

Matches 210; Conservative 73; Mismatches 107; Indels 12; Gaps 2;

QY 17 MMLLIIVAMFORNVSVHMPTRKAVDPEAFNISEIIIOHGYPCOEYEVATEDGYILSVNR 69

Db 1 MMLLIIVAMFORNVSVHMPTRKAVDPEAFNISEIIIOHGYPCOEYEVATEDGYILSVNR 69

QY 70 GYILSVNRIPRGLVOPKRTGSRPVVLLQHLVGASNMISNLPNNSIGFTIADAGFDVWM 129

Db 56 GYILSVNRIPRGLVOPKRTGSRPVVLLQHLVGASNMISNLPNNSIGFTIADAGFDVWM 115

QY 130 GNSRGNAMSRKIKTSLSDQEFMAFSYDEMARFDLPVAVINFILOKGOEKIYVYGSGGT 189

Db 116 GNSRGNAMSRKIKTSLSDQEFMAFSYDEMARFDLPVAVINFILOKGOEKIYVYGSGGT 175

QY 190 TNGFIATSPPELAIKIKMFALAPITVYKHAASPGKFFLLPDMIKIGFGKKEFLYQ 249

Db 176 TNGFIATSPPELAIKIKMFALAPITVYKHAASPGKFFLLPDMIKIGFGKKEFLYQ 235

QY 250 RELRQVILDOICSNIMLLGFPNTNNMNSRASVYAAHTLAGTSYONILHMSQAVNSGE 309

Db 236 RELRQVILDOICSNIMLLGFPNTNNMNSRASVYAAHTLAGTSYONILHMSQAVNSGE 295

QY 310 QAVNSGELRAFDMSGETKMLEKNOPTPVRYRYRDMTPTAMTGGQDMLSPNEDVKML 369

Db 296 QAVNSGELRAFDMSGETKMLEKNOPTPVRYRYRDMTPTAMTGGQDMLSPNEDVKML 355

QY 370 SEVTNLIYKNIPEMAHVDPIWGLDAPHRMYNEIILHMOOE 411

Db 356 SEVTNLIYKNIPEMAHVDPIWGLDAPHRMYNEIILHMOOE 397

RESULT 6

T33198

hypothetical protein ZK6.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000

C:Accession: T33198

R:Wu, X.

submitted to the EMBL Data Library, May 1998

A:Description: The sequence of C. elegans cosmid ZK6.

A:Reference number: 221301

A:Accession: T33198

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-403 <WU>

A:Cross-references: EMBL:AF067942; PIDN:AAC17694.1; GSPDB:GN00023; CESP:ZK6.7

A:Experimental source: strain Bristol N2; clone ZK6

C:Genetics:

A:Gene: CESP:ZK6.7

A:Map position: 5

A:Introns: 29/3; 63/1; 219/3; 319/3; 365/3

C:Superfamily: triacylglycerol lipase, lingual

Query Match 33.5%; Score 752.5; DB 2; Length 403;

Best Local Similarity 41.1%; Pred. No. 1.7e-58;

Matches 168; Conservative 66; Mismatches 142; Indels 33; Gaps 12;

QY 17 MW-LIIVAVMFORNVSVHMPTRKAVDPEAFNISEIIIOHGYPCOEYEVATEDGYILSV 75

Db 1 MW-LIIVAVMFORNVSVHMPTRKAVDPEAFNISEIIIOHGYPCOEYEVATEDGYILSV 55

A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-426 <NTL>
 A:Cross-references: EMBL:281055; PIDN:CA802896.1; GSPDB:GN00022; CESP:F01G10.7
 A:Experimental source: clone F01G10
 C:Genetics:
 A:Gene: CESP:F01G10.7
 A:Map position: 4
 A:Introns: 29/3; 64/1; 186/3; 347/2
 C:Superfamily: triacylglycerol lipase, lingual

Query Match 29.2%; Score 655.5; DB 2; Length 426;
 Best Local Similarity 35.9%; Pred. No. 6.6e-50;
 Matches 140; Conservative 85; Mismatches 144; Indels 21; Gaps 9;

QY 20 LILVAVMFORNVNSVHPTKAVDEAFNISEIIIOHOGYPCPEYEATDGYILSVNRP 79
 DB 6 LFLILSPSTNL-----SLAIDECVYTPVEIGKHFGYSEVHLVPTTDEVILELHRIP 59
 QY 80 -RGLVQPKTGSRPVYLQHGVLGASWISNLPNNSLGFTLADAGFDVMKNSRGNAWS 138
 DB 60 CKQNEKDRSSKRPVFMQHLADGFSWIPNLANSAGVFADAGFDIWIANSRGTPAS 119
 QY 139 RKHTLSIDODEFAFSYDEKARFDLPVINFIILOKGOEKIYVGYSGTGMFIARST 198
 DB 120 OKHIGYGPENOKFWNFVMOQSEFDLTASVDLVLEKOEFLYVYGHSGTGMFISRLAE 179
 QY 199 MPDLAKIKMFAFALPIATVNAKSP---GTFELLDPMMIKGLFGKEF--LYQTRFL 252
 DB 180 NPEFAKIRHHALAPATVSHIGLFGTFKTLTYAEI---LLGLPSPSPISPTTV 235
 QY 253 RQVLYVLCGOVILDOICSNIMLLGFEFNTNNNSRASVYAHTLAGTSVONILHMSQAV 312
 DB 236 OKMISYMSRFPMONICTLIDGIFDG-NEKMFNOSRGVYVLCPTPAATSKVLOHMIDLY 294
 QY 313 NSGLRAFDMSEETKNEKNOPTPVRRVDMVPTAMTGGOMLNPEVYK-MLSE 371
 DB 295 KSGVNSKFDYTD-GNILEYGOPTPEPDELQINPTVLYWSRDDILDADTDIDRSILSK 353
 QY 372 VTNLIYK-NIPEWAVDFIWLDPAPHRMY 400
 DB 354 MNKTIAGSLERPHYSHMDPFGTAAAFDLY 383

RESULT 10
 JT0949
 egg-specific protein - silkworm
 C:Species: Bombyx mori (silkworm)
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 26-Feb-1998
 C:Accession: JT0949; A28527
 R:Sato, Y.; Yamashita, O.
 submitted to JIPID, September 1991
 A:Reference number: JT0949
 A:Accession: JT0949
 A:Molecule type: DNA
 A:Residues: 1-559 <SAT>
 A:Experimental source: larva
 A:Note: this protein is a homotrimer
 R:Indraisth, L.S.; Sasaki, T.; Yamashita, O.
 J. Biol. Chem. 263, 1045-1051, 1988
 A:Title: A unique protease responsible for selective degradation of a yolk protein in Bc
 A:Reference number: A28527; MUID:88087166
 A:Accession: A28527
 A:Molecule type: protein
 A:Residues: 19-40;133-144, 'T', 146-152;229-248 <IND>
 C:Genetics:
 A:Map position: 19
 C:Keywords: egg yolk; homotrimer
 F:132-133/cleavage site: Lys-Aan (egg 30.5K cysteine proteinase) #status experimental
 F:228-229/cleavage site: Arg-Asp (egg 30.5K cysteine proteinase) #status experimental

Query Match 25.6%; Score 574.5; DB 2; Length 559;
 Best Local Similarity 33.2%; Pred. No. 1.4e-42;
 Matches 132; Conservative 82; Mismatches 152; Indels 31; Gaps 9;

QY 32 NSVH---MPTKAVDEAF-NMISEIIIOHOGYPCPEYEATDGYILSVNRPGLVQPK 87
 DB 171 NTFHDAISETORENNEDHNLNATELTKHQYVPEHFTVATDGHVTLRLRP-PTHQRD 229
 QY 88 TGSRPVYLQHGVLGASWISNLPNNSLGFTLADAGFDVMKNSRGNAWSRKHTLSID 147
 DB 230 DKRRVALLMGLGSDADWLMGSKSLAYMLDAGDVMKNSRGNAWSRKHTLSID 289
 QY 148 ODEFWAFSYDEKARFDLPVINFIILOKGOEKIYVGYSGTGMFIARSTAPLAKIK 207
 DB 290 LNFDFKFSNDELALHDLPAITIDHVLDSIGQERLHIGHSQATTFALMSQSPNEXIV 349
 QY 208 MYFALPIATVNAKSP---GTFELLDPMMIKGLFGKEF--LYQTRFLQVLYL 259
 DB 350 SMHALSPVYMNYSPLFRMIAPTSKPYOYIHQVGHGAREPGKHLIET-----F 400
 QY 260 CGQVILDO-----ICSNIMLLGFEFNTNNNSRASVYAHTLAGTSVONILHMSQAVNS 314
 DB 401 GGAACREKLRGHCNNLNIVYISGINYNODADIVPVMAHLPACTSARVKKOQONAS 460
 QY 315 GELRAFDMSEETKNEKNOPTPVRRVDMVPTAMTGGOMLNPEVYKMLSEVTN 374
 DB 461 HFRKRYNGAET-NKRYVGASEPESYDSKVASPNLYSHDAMLAHFKVDKLEENLPN 519
 QY 375 LIYKNIPEWMAH---VDFIWLDPAPHRMYNEIHLMO 408
 DB 520 VKOSFEVEQGHFDLDPQFSKKAPDYQYKLMENMQ 556

RESULT 11
 T39540
 triglyceride lipase-cholesterol esterase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
 C:Accession: T39540
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.
 submitted to the EMBL Data Library, February 1998
 A:Reference number: 221862
 A:Accession: T39540
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-443 <WOO>
 A:Cross-references: EMBL:AL021748; PIDN:CA16863.1; GSPDB:GN00067; SPDB:SPBC16A3.12C
 A:Experimental source: strain 972h-; cosmid c16A3
 C:Genetics:
 A:Gene: SPDB:SPBC16A3.12C
 A:Map position: 2
 A:Introns: 11/2; 152/3; 282/1; 398/3
 C:Superfamily: triacylglycerol lipase, lingual

Query Match 22.5%; Score 505; DB 2; Length 443;
 Best Local Similarity 29.6%; Pred. No. 1.3e-36;
 Matches 125; Conservative 91; Mismatches 154; Indels 52; Gaps 13;

QY 9 WIVSHRMEMW--LILVAVMFORNVNSVHPTKAVDEAFNISEIIIOHOGYPCPEYEYA 66
 DB 35 WIPRVEVMEFRVRLITVLYHYSKTTDGMTDA--OKCRNIYITCAFGYRVEHLYR 92
 QY 67 TEDGYILSVNRPGLVQPKT-GSRPVYLQHGVLGASWIS-NLPNNSLGFTLADAG 124
 DB 93 TQDNFILLHRI---THPKSQHREYVYCHGLMTNSELMVAANESERSLPFLVLSG 148
 QY 125 EDVMKNSRGNAWSRKHTLSIDODEFAFSYDEKARFDLPVINFIILOKGOEKIYVGY 184
 DB 149 YDVMKNSRGNAWSRKHTLYPKKDEFWNFSLDDWAMPDIDPQVYILRETRKELNLYG 208
 QY 165 YSOGTGMFIARSTAPLAKIKMYFALPIATVNAKSPGKFL--PDMIKGLFGK 242

DB 209 FSGCTAQAAMALSLINDLNDKVNIFGLAPATAPKGSNTFVUYIKVANKMYH-LFGR 267
 QY 243 KEFLYOTRFLRQLYVILGCVIJDQICSNIMLLGFFNTNNMMSRASYAATLGTATSV 302
 DB 268 RCLLPVTFWQNN---ICPPPIPVKLVDSKTLFPMWDLNSLNKICGATLHLYPSSV 323
 QY 303 QNLIHNSQAVNSGELAFD-----WGSE-----TKNLEKCNOPTRVRYRVDMT 346
 DB 324 KSVYVHMLQIINKCTFQLYDDMALLAGYSGRHYQVPLFPNNI-KC----- 368
 QY 347 VPTAMTGGODMLSNPEDEYKMLSEVTNLIYHNKINPEMAHVDIWLGLDAPHRYNLIHL 406
 DB 369 -PULLIMGKQDILINMEVMTALPPRAKEY---SIATHELDLWGDQVKEEYFPVVIDA 424
 QY 407 MQ 408
 DB 425 LK 426

RESULT 12
 T31611
 hypothetical protein Y50E8A.g - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T31611
 R:Steward, C.
 submitted to the EMBL Data Library, September 1999
 A:Reference number: 221047
 A:Accession: T31611
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1585 <MIL>
 A:Cross-References: EMBL:AL117200; NID:e1549770; PIDN:CA855050.1; CESP:Y50E8A.g
 A:Experimental source: clone Y50E8A
 C:Genetics:
 A:Gene: CESP:Y50E8A.g
 A:introns: 25/3; 60/1; 133/2; 217/3; 270/3; 337/2; 400/1; 746/2

Query Match 21.3%; Score 478.5; DB 2; Length 1585;
 Best local similarity 30.6%; Pred. No. 1.8e-33;
 Matches 118; Conservative 67; Mismatches 172; Indels 29; Gaps 9;

QY 44 EAFMNISETIIHQGYPCSEYVATEDGYILSVNRIPGLVQPKTG--SRPVLLQHGIV 101
 DB 20 EETLDADTISHGYTVERKHVYTTDDGYVOLRIPIVGRDRIILGCSRPPVYFMAHGLE 79
 QY 102 GGASNMISLNPNSLGLFIADAGFDVWGMGNSRGNMRSRKHKITLSDODEFWAFSYDEMAR 161
 DB 80 GSYHFLNLPSSAAVITADAGFDVWLGNIKGTETGLNHTSFSTINGVNFNWSLYEHS 139
 QY 162 FDLPAVINFILOKTGOEKIYVYGSGTGMGFAFSTMPEL---AQIKIMYFALAPIAT 217
 DB 140 YDLRQIEVLDYTRHESLFYVGHSGTA---VMFARLAEADVTWMSKIRVFALGPYAG 196
 QY 218 VHAHNSPGTKFLLPDMIMKGLFGKKEFLYQTRF-----LRQVLYLCGOVILDO 267
 DB 197 FKRPMP---FLLLENYDALI---QFALDGKFGILPEVIRPAISKPADFCSSKFTF 250
 QY 268 ICSNIMLLGFFNT--NNMMSRASYAATLGTATSVQNLHNSQAVNSGELAFDMSGSET 326
 DB 251 LCSAGKVAAGIETLGOVDSRIPITLSFPSTSTLNLHMQQIFKYHELRLDGLGT-A 309
 QY 327 KNLKCNOPTRVRYRVDMTVPTAMTGGODMLSNPEDEYKMLSEV--TNLIYHNKINPEM 384
 DB 310 RNLIVAGQKADARLEIGNIIAQTILYFSKDDQITDEVVDREIIMQMGGLIESYLDLF 369
 QY 385 AHVDITWGLDAPHRYNLIHLMOQE 410
 DB 370 THDFITGLRATDEVYKPIYRIYKE 395

RESULT 13

T41053
 triglyceride lipase-cholesterol esterase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T41053
 R:Hubbert, H.; Duesterhoeft, A.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, August 1998
 A:Reference number: 221967
 A:Accession: T41053
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-467 <HIL>
 A:Cross-References: EMBL:AL031324; PIDN:CAA20447.1; GSPDB:GN00068; SPDB:SPCC1672.09
 A:Experimental source: strain 972h-; cosmid c1672
 C:Genetics:
 A:Gene: SPDB:SPCC1672.09
 A:Map position: 3
 A:introns: 12/2; 164/3; 294/1

Query Match 21.1%; Score 475; DB 2; Length 467;
 Best local similarity 31.2%; Pred. No. 6.4e-34;
 Matches 125; Conservative 68; Mismatches 158; Indels 50; Gaps 12;

QY 39 KAVDPEAFMNISETIIHQGYPCSEYVATEDGYILSVNRIP-----GLVQPKTGSAP 92
 DB 73 KATHDAA--DIRMCKISGYVEDHLVTRTEDYILCIHRISKDSPGIGSPHPKTL---P 127
 QY 93 VYLLQGLVGGASNMISNL--PNNSLGLFIADAGFDVWGMGNSRGNMRSRKHKITLSDODEF 151
 DB 128 VVYCHGHLNNSVYVNCNDPRNCLVFDLVNKKYDVLGNRRNKKSRQHLREDSTDKEF 187
 QY 152 WAFSYDEMARFDPFAVINFILOKTGOEKIYVYGSGTGMGFAFSTMPELAQIKIMYFA 211
 DB 188 WDFSIDPFAQYDIPDTIDYILKTSQTKLYIGFSQGTQAPASLSIRHPLNDKINSILA 247
 QY 212 LAPATVAKHAKSPGTFKLLPDMIMG-----LEKKEFLYQTRFLRQVLYLCGOVI 264
 DB 248 LAP-----AISPRLGNHNVDAFVAKRPSILFFLEGRKSIILPSAGFQMSFL----APKF 297
 QY 265 LDQICSNIMLLGFFNTNNMMSRASYAATLGTATSVQNLHNSQAVNSGELAFDMSGSET 321
 DB 298 FPAVLAVYCLSLFNMSCNISYQRLVSFAHLVSYVCLVHWFQIMSAEFMYDNDQ 357
 QY 322 WGESE--TKNLEKCNOPTRVRYRVDMTVPTAMTGGODMLSNPEDEYKMLSEVTNLIYHK 379
 DB 358 LGHDYFLKYYKAKRPT-----NNIRTPILYILWGSDSL---VDIQAMNLNLPAREVH 408
 QY 380 NIPEMAHVDFING-----LDAPHRYNLIHLMOQEETN 413
 DB 409 KVDSEHLDIMADVAKDYVPPVRLRLDIIHHPREHND 449

RESULT 14
 E64326
 Probable lysosomal acid lipase [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: E64526
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.
 M.; Koo, H.; Motilal, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, T.; Tallor
 eus, D.; Nierman, W.C.; White, O.; Eilsen, J.A.; Salzberg, S.L.; Fraser, C.M.; Ventle
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487
 A:Accession: E64526
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-344 <STO>
 A:Cross-References: GB:AE002093; NID:g4585908; PIDN:AAD25569.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g15230
 A:Map position: 2

Query Match 19.8%; Score 445; DB 2; Length 344;

Best Local Similarity 31.2%; Pred. No. 1.9e-31;
Matches 115; Conservative 55; Mismatches 124; Indels 74; Gaps 12;

```

QY 50 SEIQHOGVCEPEYATDGYILSVNRI---PRGLVQPKKSGRPVYLQHLGGAS 105
   ::::| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 34 ADLHPANTSCHEHSIQTDGYILALQVYASLQPR-----LQSGPVYLQHLGMAGD 86
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 106 NMISNLPNNSLGLFLADAGFDVWGMNSRGNAMSRKHKTLSDODEFAFSYDEMAREFLP 165
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 87 VWFILNSPKRSIGFLADHDFDVGWGVNRTGRYSYGHVTLSDPDKEKFWMDMSWQDLAMAYDLA 146
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 166 AVINFLQKTGQKRIYVYSGGTWGFIAFTSMPELQKIKMYFALPIATVYKNAKSPG 225
   : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 147 EMIOYQ-LYSINSKIFLYVHSGGTIMSFAL-TOPHVAEVEVAALCPISYDHTAP- 203
   : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 226 TFFLLPDMIMKGLFGKKEFLVQTRRLQRLVYLCGVILDOICSNIMLLGGFNTNNM 285
   : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 204 -----LVERMV-----FMH-----LDQI----- 216
   : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 286 MSRAVYAHTLAGTSVONILHMSQAVNSGELAFPMGSETEKMLEKONOPTPVRYVRDM 345
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 217 -----EYLLDYEPHPSVKNIHRLFQMIKRGTFQYDYG-YFKRLRYGSLKPEPFIISHI 271
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 346 TVPTAMNT--GGQDWLSNPDVYKMLSEVTN--LIYKNIPEMAHVDFTGLDAPHRMY 400
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 272 PASLPWMVGSGTDLADVTDEHTLAEPLSPSELLY---LEDYGHIDFVLSAKEDVY 328
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 401 NEIHLMO 408
   : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 329 KHMIOFFR 336
   : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

RESULT 15

S37965

Probable triacylglycerol lipase (EC 3.1.1.3) - yeast (*Saccharomyces cerevisiae*)

N:Alternate names: protein YKL140W

C:Species: *Saccharomyces cerevisiae*

C:Date: 03-May-1994 #sequence:revision 03-May-1994 #text_change 06-Feb-1998

C:Accession: S37965; S39000; S22278

R:Ramezani Rad, M.; Xu, G.; Kirchbach, L.; Filtz, C.; Keuchel, H.; Hollenberg, C.P. submitted to the Protein Sequence Database, March 1994

A:Reference number: S37953

A:Accession: S37969

A:Molecule type: DNA

A:Residues: 1-548 <RAM>

A:Cross-references: EMBL:Z28140; NID:9486238; PID:9486239; MIPS:YKL140W

A:Experimental source: strain S288C

R:Abraham, P.R.

submitted to the EMBL Data Library, August 1993

A:Reference number: S38996

A:Accession: S39000

A:Molecule type: DNA

A:Residues: 1-548 <ABR>

A:Cross-references: EMBL:Z25464; NID:9396434; PID:9396439

R:Abraham, P.R.; Mulder, A.; van't Riet, J.; Planta, R.J.; Raue, H.A. yeast 8, 227-238, 1992

A:Title: Molecular cloning and physical analysis of an 8.2 kb segment of chromosome XI

A:Reference number: S22274; MUID:92245761

A:Accession: S22278

A:Molecule type: DNA

A:Residues: 1-398 <AB2>

A:Cross-references: GB:S97962

C:Genetics:

A:Gene: SGD:TGL1

A:Cross-references: SGD:S0001623; MIPS:YKL140W

A:Map position: 111

C:Keywords: carboxylic ester hydrolase

Query Match

18.8%; Score 422.5; DB 2; Length 548;

Best Local Similarity 31.6%; Pred. No. 3.5e-29;
Matches 119; Conservative 66; Mismatches 127; Indels 65; Gaps 15;

```

QY 61 EEEYVATEDGYILSVNRIIPGLVQP--KKGSRPVYLQHLGGASNMISNLP--NSISG 117
   : | : | | | | | | | | | | | | | | | | | | | | | | | | |
Db 78 EDLIVRTEDNYILTLHRIP-----PISKRNFNNKVVYLHGHGLMCDVWCNIERRKNLP 132
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 118 FLIADAGFDVWGMNSRGNAMSRKHKTLSDODEFAFSYDEMAREFLPAVINFLQKTQ 177
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 133 FVLHDLGYDVMWGMNNGNRYSTAHLNKPPKSNKFWDFSIDFAFDIPNSIEFLDITVY 192
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 178 EKIVYVYSGGTWGFIAFTSMPELQKIKMYFALPIATVK--HAK-----SPGTR 227
   : | : | : | : | | | | | | | | | | | | | | | | | | | | |
Db 193 DKYICIGFSQGSQMPAASLSSEKLNKRYSHFLAIPATPKGLNHRIVDTLAKSPGIM 252
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 228 FLLPDMIMKGLFGKKEFLVQTRRLQRLVYLCGVILDOICSNIMLLGGFNTNN 283
   : | : | : | : | | | | | | | | | | | | | | | | | | | | |
Db 253 YL-----FEGRKIVLPSAVIMQRTLHPLEFNLG-----ID-IANKIL-----FNKMS 293
   : | : | : | : | | | | | | | | | | | | | | | | | | : |
QY 284 MNM-----SRASVYAHTLAGTSVONILHMSQAVNSGELAFD-----WGSEYKMLEK 331
   : | : | : | : | | | | | | | | | | | | | | | | | | : |
Db 294 FNILPPOKIASYAKLYST-----TSVKSIYHWFQILRSQKFWQFEESDMNLSLTPYQI 348
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 332 CNOPTVRYRVNMTVPTAMTGGQDWLSNPDVYKMLSEVTNLIYKNIPEMAHVDFTM 391
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 349 ANPPTPT-----NIKIPILIIYIGIDSLVDIVAKKNLP--FNSYVDVAVDNTEHLDLIM 401
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 392 GLDAPHRMYNEIHLMO 408
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 402 GRDADTLVAKVLFIE 418
   | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: January 30, 2002, 23:02:30
Job time: 2911 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2002, 23:01:24 ; Search time 50.02 Seconds
(without alignments)
310.060 Million cell updates/sec

Title: US-09-578-063-47

Perfect score: 2247
Sequence: 1 METLSSRQNVSHRMEMILL.....IILMOQETNLSGCEAVL 423

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1289	57.4	399 1	LICH_HUMAN
2	1164.5	51.8	395 1	LIPG_RAT
3	1154	51.4	398 1	LIPG_CANFA
4	1139	50.7	397 1	LIPG_BOVIN
5	1119	49.8	398 1	LIPG_HUMAN
6	663	29.5	394 1	LIP3_DROME
7	546	24.3	439 1	LIP1_DROME
8	422.5	18.8	548 1	TGIL1_YEAST
9	126	5.6	987 1	YD94_METUA
10	105	4.7	430 1	ORC2_CAEEL
11	101	4.5	436 1	PAFA_CAVPO
12	96.5	4.3	287 1	Y193_HAEN
13	95.5	4.3	1259 1	LIN1_HUMAN
14	93.5	4.2	278 1	PRXC_STRAU
15	93	4.1	569 1	PRD_PLAFA
16	92	4.1	748 1	PTIP_ECOLI
17	91	4.0	91	P2C2_SCHPO
18	90.5	4.0	1037 1	ACRD_ECOLI
19	90	4.0	277 1	BPA2_STRAU
20	90	4.0	1122 1	YGC3_YEAST
21	90	4.0	1136 1	POLG_OMV
22	89	4.0	330 1	EXOA_RHME
23	88.5	3.9	185 1	YC2H_BACSU
24	88.5	3.9	185 1	CFAC_ECOLI
25	87	3.8	473 1	SCRB_LACIA
26	86.5	3.8	382 1	YDHH_HAEN
27	86.5	3.8	665 1	INVA_SALTY
28	86.5	3.8	1882 1	Y468_MYCPN
29	86	3.8	525 1	ACUR_NEDRN
30	86	3.8	1866 1	YGNB_CPMV
31	85.5	3.8	428 1	Y6E3_SCHPO
32	85	3.8	429 1	ARB2_ECOLI
33	85	3.8	837 1	DBPA_DICDI

34	85	3.8	4128 1	PRKD_MOUSE	P97313 mus musculus
35	84.5	3.8	875 1	AMD2_XENLA	P12890 xenopus lae
36	84	3.7	451 1	YB27_YEAST	P38293 saccharomyc
37	84	3.7	804 1	SYL_BACSU	P36430 bacillus su
38	83.5	3.7	350 1	E13B_PRUPR	P52408 prunus pers
39	83.5	3.7	375 1	NDR3_MOUSE	O94Y9 mus musculus
40	83.5	3.7	389 1	AAT_PYRAD	O94012 pyrococcus
41	83.5	3.7	456 1	YF95_YEAST	O02891 saccharomyc
42	83.5	3.7	520 1	C11A_BOVIN	P00189 bos taurus
43	83.5	3.7	520 1	C11A_PIG	P10612 sus scrofa
44	83.5	3.7	690 1	VGLH_HSVJ	P52353 human herpes
45	83	3.7	266 1	THID_SALTY	P55882 salmonella

ALIGNMENTS

RESULT	ID	STANDARD	PRT	399 AA.
1	LICH_HUMAN			
AC	P38571			
DR	01-OCT-1994 (Rel. 30, Created)			
DR	01-OCT-1994 (Rel. 30, Last sequence update)			
DR	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	LYSOSOMAL ACID LIPASE/CHOLESTERYL ESTER HYDROLASE PRECURSOR			
DE	(EC 3.1.1.13) (LAL) (ACID CHOLESTERYL ESTER HYDROLASE) (STEROL			
DE	ESTERASE) (LIPASE A) (CHOLESTERYL ESTERASE).			
GN	LIPA.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBT_TaxID=9606;			
OX	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 196-212; 277-297 AND 305-315.			
RX	MEDLINE=92042192; PubMed=1718995;			
RA	Anderson R.A., Sando G.N.;			
RA	"Cloning and expression of cDNA encoding human lysosomal acid			
RT	lipase/cholesteryl ester hydrolase. Similarities to gastric and			
RT	lingual lipases.";			
RL	J. Biol. Chem. 266:22479-22484(1991).			
RM	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISUE=Liver;			
RX	MEDLINE=94155897; PubMed=8112342;			
RA	Amels D., Merkel M., Eckerskorn C., Greten H.;			
RT	"Purification, characterization and molecular cloning of human			
RT	hepatic lysosomal acid lipase.";			
RL	Eur. J. Biochem. 219:905-914(1994).			
RM	[3]			
RP	VARIANT CSD/MD PRO-200.			
RX	MEDLINE=94195814; PubMed=8146180;			
RA	Anderson R.A., Byrum R.S., Coates P.M., Sando G.N.;			
RT	"Mutations at the lysosomal acid cholesteryl ester hydrolase gene			
RT	locus in Wolman disease.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 91:2718-2722(1994).			
RM	[4]			
RP	VARIANTS CESD ARG-129 AND PRO-129.			
RX	MEDLINE=98295576; PubMed=9633819;			
RA	Ries S., Buechler C., Schindler G.,			
RA	Jung N., Schambach A., Fehrlinger P., Vanter M.T., Belli D.C.,			
RA	Greten H., Schmitz G.;			
RT	"Different missense mutations in histidine-108 of lysosomal acid			
RT	lipase cause cholesteryl ester storage disease in unrelated compound			
RT	heterozygous and hemizygous individuals.";			
RL	Hum. Mutat. 12:44-51(1998).			
CC	-1- FUNCTION: CRUCIAL FOR THE INTRACELLULAR HYDROLYSIS OF CHOLESTERYL			
CC	ESTERS AND TRIGLYCERIDES THAT HAVE BEEN INTERNALIZED VIA RECEPTOR-			
CC	MEDIATED ENOCYTOSIS OF LIPOPROTEIN PARTICLES. IMPORTANT IN			
CC	MEDIATING THE EFFECT OF LDL (LOW DENSITY LIPOPROTEIN) UPTAKE ON			
CC	SUPPRESSION OF HYDROXYMETHYLGUTARYL-COA REDUCTASE AND ACTIVATION			
CC	OF ENDOGENOUS CELLULAR CHOLESTERYL ESTER FORMATION.			
CC	-1- CATALYTIC ACTIVITY: A STERYL ESTER + H(2)O -> A STEROL + A FATTY			
CC	ACID.			

CC -1- SUBCELLULAR LOCATION: LYSOSOMAL.
 CC -1- DISEASE: DEFECTS IN LIPA ARE THE CAUSE OF THE SEVERE INFANTILE-
 CC ONSET WOMAN DISEASE (WD) AND THE MILDER LATE-ONSET CHOLESTERYL
 CC ESTER STORAGE DISEASE (CESD).
 CC -1- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
 CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M74775; AA59519.1; -
 CC EMBL: 004285; AAB60327.1; JOINED.
 CC EMBL: 004286; AAB60327.1; JOINED.
 CC EMBL: 004287; AAB60327.1; JOINED.
 CC EMBL: 004288; AAB60327.1; JOINED.
 CC EMBL: 004290; AAB60327.1; JOINED.
 CC EMBL: 004291; AAB60327.1; JOINED.
 CC EMBL: 004292; AAB60327.1; JOINED.
 CC EMBL: 004293; AAB60327.1; JOINED.
 CC EMBL: X76488; CA54026.1; -
 CC PIR: A39315; A39315.
 CC MIM: 278000; -
 CC InterPro: IPR000073; Abhydrolase.
 CC InterPro: IPR000379; Est_lip_thioest_actsite.
 CC InterPro: IPR000734; Lipase.
 CC Pfam: PF00561; abhydrolase.1.
 CC PROSITE: PS00120; LIPASE_SER; 1.
 CC KMW: Hydrolyase; Lipid degradation; Glycoprotein; Signal; Lysosome;
 CC Disease mutation; Polymorphism.
 CC SIGNAL 1 21
 CC CHAIN 22 399
 CC LYSOSOMAL ACID LIPASE/CHOLESTERYL ESTER
 CC
 CC ACT_SITE 174 174
 CC ACT_SITE 374 374
 CC CARBOHYD 36 36
 CC CARBOHYD 72 72
 CC CARBOHYD 101 101
 CC CARBOHYD 161 161
 CC CARBOHYD 273 273
 CC CARBOHYD 321 321
 CC CARBOHYD 321 321
 CC VARIANT 16
 CC VARIANT 129 129
 CC VARIANT 129 129
 CC VARIANT 129 129
 CC VARIANT 200 200
 CC VARIANT 200 200
 CC SEQUENCE 399 AA; 45415 MW; 55F27391306B609A CRC64;
 Query Match 57.4%; Score 1289; DB 1; Length 399;
 Best Local Similarity 63.5%; Pred. No. 3.8e-104;
 Matches 235; Conservative 55; Mismatches 80; Indels 0; Gaps 0;

DB 208 FCTSMALGRPLDLIDLDGDKREFLPQSAFLKGLHCHVTHLKLGLCGLGCGF 267
 QY 280 NTNNNSRASYAAHTLAGTSVONILHMSQAVNSGELRPAFDMGSETNLEKQNPYR 339
 DB 268 NERNLMSRVDTYTHSPASVSQMLHMSQAVKFKFOARDWGSAAKFFHYNSYPT 327
 QY 340 YRVRMTPTAMTGGOMLSPEDEVKMLSEVTNLIHKINPEAHADFTMGDAPIRM 399
 DB 328 YNKKMLVPTAVMSGHDMLADYVDVNLITQITLVHESIPWEHLDFITGDAWRL 387
 QY 400 YNEIHLMOQ 409
 DB 388 YNKIINLMRK 397
 RESULT 2
 LIPG_RAT
 ID LIPG_RAT STANDARD: PRT: 395 AA.
 AC P04634;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TRIACYLGLYCEROL LIPASE, LINGUAL PRECURSOR (EC 3.1.1.3) (LINGUAL
 DE LIPASE).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-85215587; PubMed-3839077;
 RA Docherty A.J.P., Bodmer M.W., Angai S., Verger R., Riviere C.,
 RA Lowe P.A., Lyons A., Entage J.S., Harris T.J.R.;
 RT "Molecular cloning and nucleotide sequence of rat lingual lipase
 RT cDNA".
 RL Nucleic Acids Res. 13:1891-1903(1985).
 CC -1- CATALYTIC ACTIVITY: TRIACYLGLYCEROL + H(2)O = DIACYLGLYCEROL +
 CC A FATTY ACID ANION.
 CC -1- MISCELLANEOUS: LINGUAL LIPASE IS SECRETED BY THE SEROUS (YON
 CC EBER(S) GLANDS AT THE BACK OF THE RAT TONGUE.
 CC -1- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
 CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X02309; CA26479.1; -
 CC EMBL: A01157; CA000136.1; -
 CC PIR: A23045; LIPRT.
 CC InterPro: IPR000073; Abhydrolase.
 CC InterPro: IPR000379; Est_lip_thioest_actsite.
 CC InterPro: IPR000734; Lipase.
 CC Pfam: PF00561; abhydrolase.1.
 CC PROSITE: PS00120; LIPASE_SER; 1.
 CC KMW: Hydrolyase; Lipid degradation; Glycoprotein; Signal.
 CC SIGNAL 1 18
 CC CHAIN 19 395
 CC ACT_SITE 171 171
 CC ACT_SITE 371 371
 CC CARBOHYD 33 33
 CC CARBOHYD 68 68
 CC CARBOHYD 98 98
 CC CARBOHYD 270 270
 CC SEQUENCE 395 AA; 44588 MW; E601854A923522BA CRC64;
 Query Match 51.8%; Score 1164.5; DB 1; Length 395;

Query Match	Best Local Similarity	51.4%; Score 1154; DB 1; Length 398;
Matches 216; Conservative 71; Mismatches 99; Indels 12; Gaps 5;		
QY 17 MWLLTLVAVMFORNVSVH-----MPKAVDPEAFNNISEITIOHGYPCEIEYATEDG 70	: : : : : : : :	
DB 1 MWLLTLTASVIS-TLGTHTGFLGKLHP- --NPEVTMNSIQMTYWGYPAEYEYVTEDEG 56	: : : : : : : :	
QY 71 YLTASNRIIPRGVQPKKNGSRVYLLDGHGVSAGSNWISNLPNNSLGLFIADAGFDVWVG 130	: : : : : :	
DB 57 YTLGIDRLPYGKNSENIGRRVAFLOHGLASATNWNISNLPNNSLAFIADAGFDVWVG 116	: : : : : :	
QY 131 NSRGANSMRKHNTLSIIDDEFWAFSYSEDAKRPDLPAVINFILOKQGOEKIYVYGSGTT 190	: : : : : :	
DB 117 NSRGTMARRNLYSPDSYEFWAFSFDKAKDLPATIDFLIKTGQDQKLVHVGSGGTT 176	: : : : : :	
QY 191 MGFIASFMPBELAQIKMYFALAPATYAKVHAKSPGTCFLLPDMNIGKLFGEKFEPLYOTR 250	: : : : : :	
DB 177 IGFIAFSTPMPLAKRIKTFYALAPATYAKVYETTLNLTNKKLVSPFLKFLGKKIR-YPHH 235	: : : : : :	
QY 251 FLKQ-LVYILCGQVLLDIOCSNIMLLGGFNNNMNMSASYYAHTLAGTSVQNIILHS 309	: : : : : :	
DB 236 FPDQFLATEVCSRETFVLDLCSNALPFLIGCFEDTMNLSMDLYLSHPRGTSVQNVLHS 295	: : : : : :	
QY 310 QAVNGSGELRAFPWGSSETKLEKCNQPTPVRYVRVDRPTVATMGSGODLSPNEDVKML 369	: : : : : :	
DB 296 QAVKSGKFOAFPMGSPVQNMHMYHQSMPYYLJLDMHPIAVMNGNDLADPHVDLL 355	: : : : : :	
QY 370 SEVTNLIIYHKNIPEMAHVDFTWGLDAPHRMYNEIILM 407	: : : : : :	
DB 356 SKLPMLIYHKKIPRYNHLDFIWMADAPQAVVNEYISMM 393	: : : : : :	
RESULT 4		
LIPQ_BOVIN	STANDARD:	PRT; 397 AA.
AC 029458:		
DT 30-MAY-2000 (Rel. 39, Created)		
DT 30-MAY-2000 (Rel. 39, Last sequence update)		
DT 30-MAY-2000 (Rel. 39, Last annotation update)		
DE TRIACYLGLYCEROL LIPASE, PREGASTRIC PRECISOR (BC 3.1.1.3) (PREGASTRIC LIPASE) (GL) (PREGASTRIC ESTERASE) (PGE).		
GN LIP.		
OS Bos taurus (Bovine).		
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
OC Bovidae; Bovinae; Bos.		
OX NCBI_Taxid=9913;		
RM [1]		
RP SEQUENCE FROM N.A.		
RC TISSUE=Tongue serous gland;		
RA MEDLINE=95011625; PubMed=7926811;		
RA Timmermans M.Y.J., Kuipers L.P., Teuchy H.;		

DE LIPASE (GL).
GN LIPE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP
RX SEQUENCE FROM N.A.
MEDLINE=87299724; PubMed=3304425;
RA Bomder M.W., Angel S., Yarranton G.T., Harris T.J.R., Lyons A.,
RA King D.J., Pieroni G., Riviere C., Verger R., Lowe P.A.;
RT "Molecular cloning of a human gastric lipase and expression of the
RT enzyme in yeast."
RL Biochim. Biophys. Acta 909:237-244(1987).
RN [2]
RP
RX SEQUENCE OF 20-45.
MEDLINE=89325292; PubMed=2753032;
RA Bernbeek S., Blaeberg L.;
RT "Human gastric lipase. The N-terminal tetrapeptide is essential for
RT lipid binding and lipase activity."
RL Eur. J. Biochem. 182:495-499(1989).
-1- CATALYTIC ACTIVITY: TRIACYLGLYCEROL + H(2)O = DIACYLGLYCEROL +
A FATTY ACID ANION.
CC
CC -1- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC -----
CC entries requires a license agreement (see http://www.isd.sdu.ch/announce/).
CC or send an email to license@isd.sdu.ch.
CC -----
DR EMBL; X05997; CA29413.1; -
DR EMBL; X05997; CA29414.1; ALF_INIT.
DR EMBL; A01046; CAA00125.1; -
DR EMBL; A12714; CAA01053.1; -
DR PIR; S04942; S04942.
DR PIR; S07145; S07145.
DR MIM; 601980; -
DR InterPro; IPR000073; Abhydrolase.
DR InterPro; IPR000379; Est_lip_thioest_actsite.
DR InterPro; IPR000734; Lipase.
DR Pfam; PF00561; abhydrolase_1.
DR PROSITE; PS00120; LIPIASE_SER; 1.
KW Hydrolyase; Lipid degradation; Glycoprotein; signal.
FT SIGNAL 1 19
FT CHAIN 20 398 TRACYGLYCEROL LIPIASE, GASTRIC.
FT ACT_SITE 172 172 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 372 372 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 99 99 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 398 AA; 45237 MM; CD3BE1621C014F0F CRC64;
Query Match 49.8%; Score 1119; DB 1; Length 398;
Best Local Similarity 52.2%; Pred. No. 2e-89;
Matches 210; Conservative 73; Mismatches 107; Indels 12; Gaps 2;
QY 17 MLLILVAYVFNQRVNSVYHMFTRKAV-----DPEAFNISEITIOHOGYPCPEEYEVATED 69
DB 1 MMLLLTVAASLT-----SVLGTTGGLFGKLLHFGSPETVMNINISQMITYWGYNPEEYEVATED 55
QY 70 GYLIVSRIRPGVLQPKKTGSRPVVLLQHLGYGASAMNINSLPNNSIGLITADAGPVM 129
DB 56 GYLIVSRIRPGKKNKSGTGRPVVFLQHLGLASATMISLPPNNSLAFITADAGIVML 115
QY 130 GNSRGNAMSRKHKTLSDIDDEFVAFSYDENARFDLPVAVINFLQKTGOEKIYYVYGSGT 189
DB 116 GNSGNVARRNRLYYSPDSVEYFAFSPDEAKAYDLPPTIDPIYKTKOKOLHYGHSGGT 175

```

QY 190 TWMFIATPMPBLAOKIKMYFALPATATYKHAASPGTKFLLPDMKIGLFGKKEFLYGT 249
 176 TGFIFSTNPSTLAKIKFIFALAPATYKYSKLINKLRFYQSLFKFIFGKITYPHN 235
 QY 250 RFLROLVYLCGVILDOJCSNIMLLGGFNTNNMNSRASVYAATLGTSTVONILHMS 309
 236 FFDQFLATEVSCREMLNLCNSMLFTICGDSKNFTSRDYVLSHNPAGTSVQNMFWHT 295
 Db 296 QAVKSGKFOAYDMGSEVQNRHMYDQSQPPYNNVTAMNVPVIAWNGCKDLADPODGLLL 355
 QY 310 QAVNSGELAFPMGSETKLEKCONOPTPYRYRVDMTVPPTAMTGGODMLSPEDYKML 369
 296 QAVKSGKFOAYDMGSEVQNRHMYDQSQPPYNNVTAMNVPVIAWNGCKDLADPODGLLL 355
 QY 370 SEVTNLIYKKNIPEMAHVDPIGCLDAPHNMVNEIILMOOE 411
 Db 356 PKLPNLIYKKEIPFNHLDPIWAMDAPQEVYNDIVSMISEDK 397

RESULT 6
 LIP3_DROME STANDARD; PRT; 394 AA.
 AC 046108;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE LIPASE 3 PRECURSOR (EC 3.1.1.).
 GN LIP3 OR CG8823.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CANTON-S;
 RX MEDLINE=98227315; PubMed=9566193;
 RT "The Drosophila melanogaster lipase homologs: a gene family with
 tissue and developmental specific expression."
 RL J. Mol. Biol. 276:877-885(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrial J.F., Adwayani A., An H.-J., Andrews-Planckoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Bence P.V., Bertman B.P., Bhandari D., Bolshakov S.,
 RA Borova K., Botchan M.R., Bouck J., Brokstein P., Bottier P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson S.M., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Gloddek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasio P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merilov G., Mishina N.V., Modarity C., Morris J., Mostrel A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).

CC -1- TISSUE SPECIFICITY: FAT BODY.
 CC -1- DEVELOPMENTAL STAGE: ONLY AT LARVAL STAGES.
 CC -1- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
 CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.). ALSO SIMILAR TO
 CC LEPIDOPTERAN EGG-SPECIFIC AND YOLK PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).

DR EMBL: Y14367; CAA74737.1; -
 DR EMBL: AE003699; AAF54935.1; -
 DR Flybase: FBgn0023495; LIP3.
 DR InterPro: IPR000073; Abhydrolase.
 DR InterPro: IPR000379; Est_Lip_thioest_actsite.
 DR InterPro: IPR000734; Lipase.
 DR Pfam: PF00561; abhydrolase; 1.
 DR PROSITE: PS00120; LIPASE_SER; 1.
 DR Hydrolyase; Lipid degradation; Signal; Glycoprotein.
 FT SIGNAL 1 20
 FT CHAIN 21 394
 FT ACT_SITE 164 164
 FT ACT_SITE 369 369
 FT CARBOHYD 131 131
 SQ SEQUENCE 394 AA; 44901 MW; A718D1D743673802 CRC64;

Query Match 29.5%; Score 663; DB 1; Length 394;
 Best Local Similarity 37.0%; Pred. No. 6; 3e-50;
 Matches 137; Conservative 80; Mismatches 139; Indels 14; Gaps 7;

QY 51 EIIHOGYPOEYEVATEDGYILSVNRIPGLVQPK--KTGSRPVYLLQHLGVGASNNI 108
 30 ERIEDDGYPMERHEVYTSIDVYILTHRIP---YSPKGTSSNRPVAFLMHGMLSSSDWV 86
 Db 109 SNLPNNSLGFILADAFDVMGSRGNMNSRRKKTISTODDERMAVSYDEMAFDLPAYI 168
 87 LMPERSLAVYMLADADYDVMGNAKNTYSKAKRWPTWQIFWFSNMEIGYDVPAMI 146
 QY 169 NFILOGTGOKIYVYVSOGTGMFIASFSTMPBLAOKIKMYFALPATATYKHAASPGTK 227
 147 DYFLAKTGQGOQVYVHSGOITYYLVWSERPYNKIKSAHLGLPAALMGNNKSLPTGA 206
 Db 228 ---FLLPDMITKGLFGKKEFLYOTRFLROLVYILGQYI-LDOJCSNIMLLGGFNTNN 283
 207 FAVILQOPNAIVE-VGSMSEFMSNFKODLIGEMCOATSPYADMCANEIFLIGDYDTEQ 265
 QY 284 MNMSRSVYAATLGTSTVONILHMSQAVNSGELAFPMGSEFKNLEKCONPTPYRYR 343
 266 LDVELLEHRIKATSPASAVNONLHFOEANSNGFRFEDY-TALRNYEYGSFFPPYK 324
 QY 344 DMVPTAMTGGODMLSPEDYKMLSEVYNTLIYHKNIP--EWAHVDPIGDLADPAPRYN 401
 325 NAKAPVLLIYGANDMCDVSDYAKLDELCPNMAALDVLVPEKRAHDFITMGTFARAYVD 384
 Db 402 EIIHMOOE 411
 385 EVLKQMSYE 394

```

RESULT 7
LIP1_DROME STANDARD; PRT; 439 AA.
ID LIP1_DROME STANDARD; PRT; 439 AA.
AC Q64107; Q64107;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE LIPASE 1 PRECURSOR (EC 3.1.1.-).
GN LIP1 OR CG7279.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyroidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE OF 7-433 FROM N.A.
RC STRAIN=CANTON-S;
RX MEDLINE=98227315; PubMed=956193;
RT "The Drosophila melanogaster lipase homologs: a gene family with
RT tissue and developmental specific expression."
RT J. Mol. Biol. 276:877-885(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amatiades P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
RA Adair J.F., Agbayani A., An H.-J., Andrews-Plankkoc C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fouts R., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,
RA Jaitani M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA Palaret K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spredling A.C., Stapleton M., Strong R., Sun E.,
RA Styksas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: COULD BE A DIGESTIVE ENZYME.
CC -1- TISSUE SPECIFICITY: OVARIES.
CC -1- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.). ALSO SIMILAR TO
CC LEPTOTERAN EGG-SPECIFIC AND YOLK PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y14366; CAAT74736.1; -
DR EMBL: AE003629; AAF52994.1; -
DR FlyBase: FBgn0024747; LIP1.
DR InterPro: IPR000073; Abhydrolase.
DR InterPro: IPR000379; Est_lip_thioest_actsite.
DR Pfam: PF00561; abhydrolase.1.
DR PROSITE: PS00120; LIPASE_SER; FALSE_NEG.
KW Hydrolase; Lipid degradation; Signal; Glycoprotein.
FT SIGNAL 1 24
FT CHAIN 25 439
FT DOMAIN 30 44
FT ACT_SITE 197 197
FT ACT_SITE 393 393
FT CARBOHYD 124 124
FT CARBOHYD 151 151
FT CARBOHYD 346 346
FT CARBOHYD 379 379
FT CARBOHYD 426 426
FT CONFLICT 10 10
FT CONFLICT 213 213
FT CONFLICT 412 412
SQ SEQUENCE 439 AA; 50660 MW; 9E32E0BEAE93BF CRC64;
Query Match 24.3%; Score 546; DB 1; Length 439;
Best Local Similarity 31.8%; Pred. No. 9,9e-40;
Matches 123; Conservative 83; Mismatches 157; Indels 24; Gaps 7;
OY 39 KAIDPEAPNMISEIIOHGCPCEYEYATPDEYILSVNRIPRGLVQPKTGSRPVYLQ 98
DB 58 KNKQDSTLSVDKLINKKYGESEVHHVTTEDGILTMHR-----RKQGA-PELLOH 109
OY 99 GLVGASNNWISLNPNSLFLIADAGFDVWNGSRGNANRKRKKTLSIDDEFMAFSYDE 158
DB 110 GLVDSAGFVWMPNVLAVIADHDHYDVLGNARNRYSRNHTTLPDPSKFWDSWBH 169
OY 159 MAFEDIPAVINFLQKTGCEKITYYGVSGCTMGFLAFSPMPRLAKIKYFLAIYAY 218
DB 170 IGWYDIPAMIDHVLKTYGPKLHYACHSOGCSFVWCSRPAYNDKYSWMLAPAYYA 229
OY 219 KHAKS-PGTRFLLPMMKIGLFGKKEFLYOTRFRLQVLYLGGVILDOICSNMILLG 277
DB 230 KETEDHPYIRALISLYNSLVGSSIRFMFNGCFELCRM-----ETERLCLEAVFGIV 283
OY 278 GFNTNNMNSRASVYAHTLAGTSVONILHWSQAVNSGELRAFDWGSSETKNLECKNQPP 337
DB 284 GRMNEFNKRMPFVILGHYPAGVAQVKHFIIQIKSGRFAFSY-SSNNKMLQYRDHP 342
OY 338 VRRRVADNVPYPMATMGQDMLSNPDYKMLSEVYNLIYHKIIP--EAAHVDFINGDA 395
DB 343 PRNLISLVYVFPFYVYSTINDLCPKPDVSMCDLGNVTKYLVPOKEFNHMDLWAIYD 402
OY 396 PHRYNEIITHLM-----OOEETNLSQ 416
DB 403 RKMILYRMLQVIGLKVGESEPEEANKSR 429
RESULT 8
TG1L YEAST STANDARD; PRT; 548 AA.
ID TG1L YEAST STANDARD; PRT; 548 AA.
AC P34163;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE TRICLIGERIDE LIPASE-CHOLESTEROL ESTERASE (EC 3.1.1.-).
GN TG1L OR YK1140W OR YK15.
OS Saccharomyces cerevisiae (Baker's yeast).

```

CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92245761; PubMed=1574929;
 RA Abraham P.R., Mulder A., Van't Riet J., Planta R.J., Rane H.A.;
 RT "Molecular cloning and physical analysis of an 8.2 kb segment of
 RT chromosome XI of Saccharomyces cerevisiae reveals five tightly linked
 RT genes."
 RL Yeast 8:227-238(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Red M.R., Xu G., Kirchhath L., Fritz C., Keuchel H., Hollenberg C.P.;
 RL Submitted (MAR-1994) to the EMBL/Genbank/DBJ databases.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: Z25464; CAAB0958.1;
 DR EMBL: Z28140; CAAB1981.1;
 DR PIR: S37969; S37969.
 DR PIR: S39000; S39000.
 DR SGD: S0001623; TGL1.
 DR InterPro: IPR000073; Abhydrolase.
 DR InterPro: IPR000379; Est_11P_Cholest_actsite.
 DR Pfam: PF00561; abhydrolase.1.
 DR HydroLase; Lipid degradation.
 KW
 SQ SEQUENCE 548 AA; 62979 MW; 32D1F230701CB083 CRC64;

Query Match 18.8%; Score 422.5; DB 1; Length 548;
 Best Local Similarity 31.6%; Pred. No. 6.7e-29;
 Matches 119; Conservative 66; Mismatches 127; Indels 65; Gaps 15;

QY 61 EEYEVATEEGYILSVNRIRGLVOP--KITGSRPVYLLDHGLVGGASNNISNP--NNSLG 117
 DB 78 EHLVTRTEEDNYILTLTRIP-----PISKRFRNNKVYLLHGLMCDVVCNTERHKNLP 132
 QY 118 FILADAGEVYMGNSRGNAMSRKHKTLSDIDEFMAFSYDEARFDPVAVINFILOKTCQ 177
 DB 133 FVLHDLGYDVWGNNGNKNYSTAHLNPKPKSNKPFDFSIDFAFPDIPRSIEFIIDITKV 192
 QY 178 EKIVYVGSQGTWGFIASTPELAQIKMTFALAPATVK--NAK-----SPGTR 227
 DB 193 DKVICTIGSOGSAOMFAFSLSEKLNKRVSHFLAIPAMTPKGLHNRIVDTLAKSSPGFM 252
 QY 228 FLLLDMMKGLFGKKEFLYOT---RFLRLVYILCGVVIDOJCSNMLLGGNTNN 283
 DB 233 YL-----FFGRKIVPSAVIQORTLHPTLFNLC---ID-IANKIL---FNMS 293
 QY 284 NNM-----SRASYAAHTLAGTSVONILHMSQAVNSGELBAFD---MGSETKLEK 331
 DB 294 FHLRLRQKIASTAKLST-----TSKSTIVHNFQILRSKPFQMFESDMKLSLTPPYOI 348
 QY 332 CNOPTPVRYRVDMTPVPAAMTGMODWLSNPEDVKMLLSEVTNLVILHKNIPMAHVDFTW 391
 DB 349 ANEPFRT-----NIKIPILLIYGIDSLVDIDVWKNKLP--FNSVEDVAVDVNHEHDLIM 401
 QY 392 GIDAPHRKATNEIHLAQ 408
 DB 402 GRDADTLVIAKVLRIE 418

RESULT 9
 ID YD94_METJA STANDARD: PRT: 987 AA.
 AC 058789;

DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL PROTEIN MJ1394.
 GN MJ1394.
 OS Methanococcus jannaschii.
 CC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 CC Methanococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.O., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kellerauge A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Belch C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geophagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hirst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Weese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii."
 RL Science 273:1058-1073(1996).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: TO M.JANNASCHII MJ1393 AND A.FULGIDUS AF2028.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U67579; AAB99404.1;
 DR TIGR: MJ1394;
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 12 32 POTENTIAL.
 FT TRANSMEM 958 978 POTENTIAL.
 SQ SEQUENCE 987 AA; 112360 MW; D1E628FB28CA86D CRC64;

Query Match 5.6%; Score 126; DB 1; Length 987;
 Best Local Similarity 23.1%; Pred. No. 0.0072;
 Matches 105; Conservative 56; Mismatches 149; Indels 144; Gaps 29;

QY 3 ETLSRQWIVSHRMKMLLIVAYMFOR-----NVNSVHPTKAVDEAFNM----- 48
 DB 340 EYDVYTKIPANKMSEYTNVLIYLNKNLVNNGENVNVNIKYLSNGQFYDNEEMTGMNQ 399
 QY 49 -----ISEIIHQG-YPCEEYEVATEDGYILSVNRIPRG--LVQPKTKGSRPVYLLQH 98
 DB 400 YKNGIYQMSISIOHSGNSLEKYGIST-----SLNDNGYKILLPKEIG-RDVIY--S 450
 QY 451 GMYVYRSNM--GGGPIDRIG--LEDENF-----GYSEVYHNSYISIDRTGNGPTE 500
 QY 149 ---DEFMAFSYDEARFDPVAVINFILOKTCQEKIYVYVSGITMGFLAESTMELAK 205
 DB 501 ISPEVYWMNPDEDEWYFEL-----KI---YSNGT---ITFST----- 531
 QY 206 IMYFALAPI-ATVYHAKSPGTRF-----LLPDMKIG---LEGKKEFLYOTR 250
 DB 532 ---YONGSLAFTVSTIDTFTYKEDRVYIHGIVYVVDLEVNSKNPFGCKANKYL-- 586
 QY 251 FLRLVYILCGVVIDOJCSNMLLGG-----FNTNMMNSRASYAAHTL-----AG 299
 DB 587 -----ETISANSEGTAVLFDGDYFKKDVNTSLN---ALNMTNITLNSNSDA 632
 QY 300 TSVONIL-HWSQAVNSGELRAFDWGSSETKNLEKQNPVPR--YRVDMTVTPTAAMTGGQ 356

DB 633 TLVFNVLGNYSYERDNLILAKYPAKILFNNGTNTSIRKGVASGSYSTDHGTCE 692
OY 357 --DWLSN---PDVAKMLSEVTNLIYHKRIPEMA 385
DB 693 INIWIENWTFKNDKAKSYFNLTNL---NI--WA 720

RESULT 10
ORC2_CAEEL STANDARD: PRT: 430 AA.
ID ORC2_CAEEL
AC Q21037;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PROBABLE ORIGIN RECOGNITION COMPLEX SUBUNIT 2.
GN F39E10.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Swaburne J.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP REVISIONS.
RA Jones S.J.M.;
RL Submitted (MAR-1997) to the SWISS-PROT data bank.
CC -1- FUNCTION: COMPONENT OF THE ORIGIN RECOGNITION COMPLEX (ORC) THAT
CC BINDS ORIGINS OF REPLICATION. IT HAS A ROLE IN BOTH CHROMOSOMAL
CC REPLICATION AND MATING TYPE TRANSCRIPTIONAL SILENCING. BINDS TO
CC THE AAS CONSENSUS SEQUENCE (ACS) OF ORIGINS OF REPLICATION IN AN
CC ATP-DEPENDENT MANNER.
CC -1- SUBUNIT: ORC IS COMPOSED OF SIX SUBUNITS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: BELONGS TO THE ORC2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; 236949; CAA85415.1;
DR WormPep; F59E10.1; CE11492.
KW DNA replication; Nuclear protein.
SQ SEQUENCE 430 AA; 49319 MM; 54FC086BB4AD9670 CRC64;

Query Match 4.7%; Score 105; DB 1; Length 430;
Best Local Similarity 21.9%; Pred. No. 0.15;
Matches 61; Conservative 42; Mismatches 87; Indels 88; Gaps 13;

OY 85 PAKTSGRPVVLQHGVLGAGSNWISLDPNN-----SLGF-----TLADAGEDVV--MGN 131
DB 25 PEKEGRS-----OKRTGKKNASRNLSNLEEDLEQLGFEDETVSMAOSAIENFMQK 78
OY 132 S-----RGNAISRKH-----TLSTIDDEFNAFSYDEMARPDLPAVINFILOKTOQEKY 181
DB 79 SASEPMNNAKSRGRRAAGNGNTELEDEDSNALTDTKCDLPGLRNYITRKKNTEFEK 138
OY 182 YVGSQGTTFMFIATSTPELAKIKMYFALAPIATVTHAKSPGKFLLLPDMITKLEFG 241
DB 139 RLEHLADNDPFG-----KKKLYLAG-----FNILHGVGS 168
OY 242 KKEPLVYOTRPLRQVLYIYAGVILIDQISNIMLLIGFNTN--NMNNSRASYAHTLA 298
DB 169 KRDLV--TEFENELSDTYMREDAKDLINVKVLLGALINEMKLNCKVNRG----- 217
OY 299 GTSVONILHWSQAV-----NSGEL-----RAFDKSGE 325

DB 218 -----OSTSMARSIRKRNNSQOLLILIDNTEAPDRSD 251
OY 218 -----OSTSMARSIRKRNNSQOLLILIDNTEAPDRSD 251

RESULT 11
PAPA_CAVPO STANDARD: PRT: 436 AA.
ID PAPA_CAVPO
AC P70683;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PLATELET-ACTIVATING FACTOR ACETYLHYDROLASE PRECURSOR (EC 3.1.1.47)
DE (PAP ACETYLHYDROLASE) (PAP 2-ACYLHYDROLASE) (IDL-ASSOCIATED
DE PHOSPHOLIPASE A2) (IDL-PLA(2)) (2-ACETYL-1-ALKYLGLYCEROPHOSPHOCHOLINE
DE ESTERASE) (1-ALKYL-2-ACETYLGLYCEROPHOSPHOCHOLINE ESTERASE).
GN PLA2G7 OR PAFAH.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HARTLEY; TISSUE=Liver;
RX MEDLINE=97103479; PubMed=8947850;
RA Karasawa K., Kuge O., Kawasaki K., Nishijima M., Nakano Y., Tomita M.,
RA Yokoyama K., Setaka M., Nojima S.;
RT "Cloning, expression and characterization of plasma
RT platelet-activating factor acetylhydrolase from guinea pig";
RL J. Biochem. 120:838-844(1996).
CC -1- FUNCTION: MODULATES THE ACTION OF PLATELET-ACTIVATING FACTOR (PAF)
CC BY HYDROLYZING THE SN-2 ESTER BOND TO YIELD THE BIOLOGICALLY
CC INACTIVE LYSO-PAF. HAS A SPECIFICITY FOR SUBSTRATES WITH A SHORT
CC RESIDUE AT THE SN-2 POSITION. IT IS INACTIVE AGAINST LONG-CHAIN
CC PHOSPHOLIPIDS.
CC -1- CATALYTIC ACTIVITY: 2-ACETYL-1-ALKYL-SN-GLYCERO-3-PHOSPHOCHOLINE +
CC H(2)O -> 1-ALKYL-SN-GLYCERO-3-PHOSPHOCHOLINE + ACETATE.
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- SIMILARITY: PARTIAL WITH OTHER LIPASES (PANCREATIC, GASTRIC,
CC HEPATIC, LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; D67037; BA11054.1;
DR InterPro; IPR000379; Est_lip_thioest_actsite.
DR InterPro; IPR000734; Lipase.
DR PROSITE; PS00120; LIPASE_SER_1.
KW Hydrolyase; Lipid degradation; Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 436
FT FT
FT ACETYLHYDROLASE.
FT ACT_SITE 271 271
FT ACT_SITE 294 294
FT ACT_SITE 349 349
FT CARBOHYD 76 76
FT CARBOHYD 200 200
FT CARBOHYD 324 324
SQ SEQUENCE 436 AA; 49062 MM; C359D96E392FE11 CRC64;

Query Match 4.5%; Score 101; DB 1; Length 436;
Best Local Similarity 20.4%; Pred. No. 0.33;
Matches 85; Conservative 60; Mismatches 149; Indels 122; Gaps 21;

OY 21 ILVAYMFORNVNSYMPKAVD-PEAFNINSEIIOHQGYPCPEY-EVATED-----GY 71
DB 70 LMSGYTNOSSFLRLRYPSQDNDPFDALM-----IPNDEYFQGLTEFTLGASSPFLGK 119

QY 72 ILSTV-----NRIPRLVOPKRTGSR-PVYLLOHGLVGASNMISLNPNSLGFILADAGFD 126
 DB 120 LKLLGVSXVPKAKWNSPLKTKGKPLIFSHG--GAFRSI-----YSALIGELASHGFI 173
 QY 127 V-----WMGNSRGNAWSKHKHKLTLSDIDE-----FW 152
 DB 174 VAAVEHDESAATYYTODAPAAESGNRSWYYKYGNLIEFERKROLROREECQSALSW 233
 QY 153 AFSDYEMARDELPAVINFILOK---TGOEKIYYVGSQGTMMGFIASTMPBLAQIKM 208
 DB 234 LLSIDEEBPKNVLDLNFIDQLGSLDRSKVAILIHSFG--GATVIQTISE-DQRFRC 289
 QY 209 YFAL---APAIATVKAHAKSPGTFLLPDMKIKGLFGKK-ELLQTRRLROLV----- 256
 DB 290 GIALDPMMFPYGEDVHSHKIPQLFFINSEYFQSANDTKIKKIFYPQKEREKMIAYKSVH 349
 QY 257 ---YLGGVILDDICSNIMLLGGFNTNNNM---SRASYAANTLAGTSYONILH 307
 DB 350 HNFVDEFTFPAFGKIT-----GQMLSLGKIDSEYAMDLINKASLFLQKYLGLD-KNFDQ 402
 QY 308 WSOAVNSGELRAFDMGSETKNLEKNOPTPVRYRYRDMTVPETAMWTGQDMLSNPE 363
 DB 403 WNSLME-----GDDENL-----IPEFTIPTMGSSGTGQRNPD 436
 RESULT 12
 Y193_HAEIN STANDARD; PRT; 287 AA.
 ID Y193_HAEIN 005013;
 AC 057427; 005013;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PUTATIVE ESTERASE/LIPASE H10193 (EC 3.1.1.-).
 GN H10193.
 OS Haemophilus influenzae
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae.
 OC Haemophilus.
 OX NCBI_Taxid=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 Fine L.D., Fritchman J.L., Fuhrmann J.L., Georgagen N.S.M.,
 Grehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 Venter J.C.;
 RA "Whole-genome random sequencing and assembly of Haemophilus
 influenzae Rd.";
 RT Science 269:496-512(1995).
 RL
 CC -1- SIMILARITY: BELONGS TO A FAMILY OF ESTERASES THAT GROUPS TOGETHER
 CC PSEUDOMONA TROPINSTERASE, DMPD, TOOF AND XYL.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: D32704; AAC21862.1; -
 DR TIGR: H10193; -
 DR InterPro: IPR000073; Abhydrolase.
 DR InterPro: IPR000379; Est_1ip_chloest_actsite.
 DR Pfam: PF00561; abhydrolase; 1.
 KW Hypothetical protein; Hydrolase; Serine esterase; Complete proteome.
 FT ACT_SITE 119 119 BY SIMILARITY.

FT ACT_SITE 266 266 BY SIMILARITY.
 SO SEQUENCE 287 AA; 32987 MW; F2B548619C060619 CRC64;
 Query Match 4.3%; Score 96.5; DB 1; Length 287;
 Best Local Similarity 19.1%; Pred. No. 0.48;
 Matches 62; Conservative 39; Mismatches 119; Indels 105; Gaps 12;
 QY 84 QPKTGSPPVLLDGLVYGASNMISLNPNSLGFILADAGFDVWNGSRGNASRKIKT 143
 DB 39 QVKQITNPVLIIFHGLGDMDN-----LGLV-----ARASEHSI 75
 QY 144 LSIODEFMAISYEMARFDLPAV-INFILOKGTGEEKIYYVGSQGTMMGFIASTMPBL 202
 DB 76 LRIDLRNGHSHFSEKKNYQLAEDVIAVIRHLNSKYLIGHSGGTAKKTALACPEL 135
 QY 203 AQIKMYFALPIATVKAHAKSPGTFLLPDM-----MIKGLFGKKEFLYQTR 250
 DB 136 VEK-----LVIDMSPPYEGFGHKDVFNGLFAVKNKAPEN- 171
 QY 251 FLRLVLIYLCQVILDDICSNIMLLGGFNTNNNMRSASYAHTLAGTSYONILHWSQ 310
 DB 172 --RQAKPILKQEIINDE--DVQVFMKSFQVNSADCFEPNL---TALFNMYANTIME- 222
 QY 311 AVNSGELRAFDMGSETKNLEKNOPTPVRYRYRDMTVPETAMWTGQDMLSNPEYKMLLS 370
 DB 223 ---KRVF-----TPTLFKGNSSYIKIENSEKILE 251
 QY 371 EVTNLIYH--KNIPMAHV---DFI 390
 DB 252 QEPNATFTINGSSGHVVAEKPDFV 276
 RESULT 13
 LIN1_HUMAN STANDARD; PRT; 1259 AA.
 ID LIN1_HUMAN 008547;
 AC P08547;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 01-AUG-1988 (Rel. 08, Last annotation update)
 DE LINE-1 REVERSE TRANSCRIPTASE-HOMOLOG.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=86230917; PubMed=2423883;
 RA Hattori M., Kuhara S., Takenaka O., Sakaki Y.;
 RT "L1 family of repetitive DNA sequences in primates may be derived
 RT from a sequence encoding a reverse transcriptase-related protein.";
 RT Nature 321:625-628(1986).
 RL
 CC -1- MISCELLANEOUS: THIS SEQUENCE WAS CONSTRUCTED FROM AN ALIGNMENT OF
 CC PUBLISHED AND UNPUBLISHED SEQUENCES, DETERMINED IN VARIOUS
 CC LABORATORIES, BELONGING TO THE LINE-1 FAMILY.
 CC
 CC PIR: A25313; GNRHL1.
 DR InterPro: IPR000097; AP-endonuclease_family-1.
 DR InterPro: IPR000300; IPPC.
 DR InterPro: IPR000477; RVTSE.
 DR Pfam: PF01260; AP-endonuclease; 1.
 DR Pfam: PF00078; rvt; 1.
 DR SMART: SM00128; IPPC; 1.
 KW RNA-directed DNA polymerase.
 SO SEQUENCE 1259 AA; 147112 MW; F3BC18A3803919F5 CRC64;
 Query Match 4.3%; Score 95.5; DB 1; Length 1259;
 Best Local Similarity 19.7%; Pred. No. 4.4; Indels 129; Gaps 23;
 Matches 90; Conservative 67; Mismatches 170; Indels 129; Gaps 23;
 QY 28 QGNVSVHPPTKAVDPPEAFMNI---SEIIHOGYGCCEYE----- 64
 DB 442 QEEVESLNRPINSSLEIALINSLPNKKSPPGPGFTAFYQRKKELVFPLKLPQSIKE 501

QY 65 -VATEDGYILSVNRIPR-GLVOPKTKGSRPVYL-----LOHGLVGCASNMISNL-PNN 114
 DB 502 GILPNSFYASILLIPKPRDPTTKENFRPISLMDARILKLNLANOIOOHKKLIHHD 561
 QY 115 SLGFIADAGFDVWGNNSR-----GNAMSKHKTLSIDODEFAFASVDEM----- 159
 DB 562 QVGFIFRANMG---WFNIRKSINICHINTKOTNMHIIISIDAEK---AFDKIQPFMLK 614
 QY 160 -----ARPDPAVINFILOKTGOEKIYVGSQGTMGFIASFTEPDELA 203
 DB 615 PLNKLGIDGTYLKIRAIYDKPTA-NIIL--NGQ--KLEAPLKTKRGCCPLSPLLPNIV 670
 QY 204 QIKIMFALAPATYVKAHSPGKFLLLPDMMIKGLFGKKEFLYOTRFLRLQVLYIYICGV 263
 DB 671 LEV-----LA--RAIRQKE-----IKGIOLKEEVEKLSLFADMDIMVYLENP 711
 QY 264 ILDOICSNIMLLGFGN-----TNMNMGRASYAAHTLAGTSYONILHMSQAVNSGELR 318
 DB 712 VSNQ---NLKLISNFSKYSYKINOKSOAFILTNRRQTESQINSLEPFTIASKRIKYL 768
 QY 319 AFDWGSETKLEKCNOPFVRYRVDMT-----VPTAMTGGQDWLSNPDVYKMLSEVT 373
 DB 769 GIGLTFDVKDLFKENY-KFLNLEIKEDTKMKKNIPCS-----WGRINIVYKMAI--LP 818
 QY 374 NLIYKKN-IP-----EMAHVDFIWDGADPH 397
 DB 819 KVIYRFNAPIKILPMTFELETETLKFIMNOKRAH 854

RESULT 14

PRXC_STRAU STANDARD; PRT; 278 AA.

AC O31168;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE NON-HEME CHLOROPEROXIDASE (EC 1.11.1.10) (CHLORIDE PEROXIDASE)
 DE (CPO-T) (CHLOROPEROXIDASE T).
 GN CPO OR CPO.
 OS Streptomyces aureofaciens.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_Taxid=1894;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-TUE24;
 RC Pelletier I., Altenbuchner J., van Pee K.-H.;
 RA Submitted (Oct-1997) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RN X-RAY CRYSTALLOGRAPHY (1.75 ANGSTROMS).
 RC STRAIN-TUE24;
 RX MEDLINE=98307994; PubMed=9642069;
 RA Hofmann B., Tolzer S., Pelletier I., Altenbuchner J., van Pee K.-H.,
 RA Hecht H.-J.;
 RT "Structural investigation of the cofactor-free chloroperoxidases";
 RL J. Mol. Biol. 279:889-900(1998).
 CC -1- CATALYTIC ACTIVITY: 2 NH + 2 CHLORIDE + H(2)O(2) -> 2 RCL + 2
 CC H(2)O.
 CC -1- SUBUNIT: HOMODIMER.
 CC -1- SIMILARITY: TO OTHER BACTERIAL NON-HEME BROMO- AND CHLORO-
 CC PEROXIDASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AF031242; AAB86626.1; -
 CC PDB; 1A7U; 17-JUN-98.

DR PDB: 1A8U; 17-JUN-98.
 DR InterPro: IPR000073; Abhydrolase.
 DR InterPro: IPR000379; Est_Lip_Thioest_actsite.
 DR Pfam: PF00561; abhydrolase; 1.
 KW Oxidoreductase; Peroxidase; 3D-structure.
 FT ACT_SITE 99 BY SIMILARITY.
 FT ACT_SITE 229 BY SIMILARITY.
 FT ACT_SITE 258 BY SIMILARITY.
 SQ SEQUENCE 278 AA; 30354 MW; 8C97A87251FBEDE5 CRC64;

Query Match 4.2%; Score 93.5; DB 1; Length 278;
 Best Local Similarity 22.6%; Pred. No. 0.84;

Matches 72; Conservative 36; Mismatches 114; Indels 97; Gaps 17;

QY 91 RPVVLLOHGLVGCASNMISNLPPNSIGFLIADAGDVWNGRGNAMSKHKTLSIDODE 150
 DB 25 QPVVLI-HQFPLSGSW-----EROSALIDAGRGVITYIDRKGFGSSQPT----- 70
 QY 151 FMAFSYDEMARDLPAVINFILOKTGOEKIYVGSQGTMGFIASFTEPDELAOKIKMYF 210
 DB 71 -GYVDYDFA-----ADNTVLETLDDDAVLGVFSMGT-----GEVARYVSSV- 112
 QY 211 ALAPATYVKAHSPGKFLLLPDMMIKGLFGKKEFLYOTRFLRLQVLYILCGVYILDQICS 270
 DB 113 GTARIKAFAFLASL-EPELTKTDNDPDGA-APKEP-----FDGIVA 151
 QY 271 NIMLLGFGFNTNMNMGRASYAAHTLAGT-----SYONILHMSQAVNSGELRA----FD 321
 DB 152 AVKADRYAFYTGFEFN---DEFYNLDENLGTRISEAVRN--SWNTAASGGFFPAAAPPT 205
 QY 322 WGESEK-NLEKCNOPFVRYRVDMTVPAMTGGQDWLSNPDVYKMLSEVTNLIYKKN 380
 DB 206 WYTFDRADIPRIDVPALLIHGTGDTLPI-----ENTAVPFKA 244
 QY 381 IP--EMAHVDFIWDGADPH 397
 DB 245 LPSAEYVEVE-----GADH 258

RESULT 15

PRXD_PLAFA STANDARD; PRT; 569 AA.

AC Q08210;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE DIHYDROROTATE DEHYDROGENASE HOMOLOG PRECURSOR (EC 1.3.3.1)
 DE (DIHYDROROTATE OXIDASE).
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_Taxid=5833;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN-3D7;
 RX MEDLINE=94049995; PubMed=8232427;
 RA Leblanc S.B., Wilson C.M.;
 RT "The dihydroxylate dehydrogenase gene homologue of Plasmodium
 RT falciparum";
 RL Mol. Biochem. Parasitol. 60:349-352(1993).
 CC -1- CATALYTIC ACTIVITY: L-DIHYDROROTATE + O(2) -> OROTATE + H(2)O(2).
 CC -1- COFACTOR: FAD.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE (PROBABLE).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

DR	EMBL: L15446; AAC37170.1; -		
DR	InterPro: IPR001295; DHO_dh		
DR	InterPro: IPR003009; FMN_enzyme		
DR	Pfam: PF01180; DHOdehase; 1.		
DR	PROSITE: PS00911; DHOdehase_1; 1.		
DR	Pyrimidine biosynthesis; Oxidoreductase; Flavoprotein; FAD;		
KW	transit peptide; Mitochondrion; Inner membrane.		
FT	TRANSIT 1 ?	MITOCHONDRION (POTENTIAL).	
FT	CHAIN ?	569	
FT	NP_BIND 235	243	
SO	SEQUENCE 569 AA; 65558 MW; 888880384EBD52FE3 CRC64;		
			DIIHDROGRATE DEHYDROGENASE
			FAD (NMD PART) (POTENTIAL).
			HOMOLOG

Query Match	4.1%	Score 93	DB 1	Length 569
Best Local Similarity	28.8%	Pred. No. 2.5		
Matches	30	Conservative	14	Mismatches 52; Indels 8; Gaps 1;

Oy 218 VKHASPSTKFLLEPDMMIKSLPKFKETLYOTRFRLDVIYLGQVLLDICSINMLLG 277
 Db 133 IKKEKSRKKRIIFLLFVSLFGLVFPFSSYNPFLLYDIFLKKCAKYIDGELCHDLFLLG 192
 Oy 278 GFNTNNMMASRASYAAHTLGTSTVQNTLHMSQAVNGELRAPD 321
 Db 193 KYNILPYDTSMDSIYAC-----TNKKHLDITNPFVAAAFD 228

Search completed: January 30, 2002, 23:10:17
Job time: 533 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2002, 23:00:04 ; Search time 91.96 Seconds
(without alignments)
672.827 Million cell updates/sec

Title: US-09-578-063-47

Perfect score: 2247
Sequence: 1 METLSRQWIVSHRHEMML.....IHLMOQETNLSGRCEAVL 423

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL.17:*
1: SP_Archaea:*
2: SP_Bacteria:*
3: SP_Fungi:*
4: SP_Human:*
5: SP_Invertebrate:*
6: SP_Mammal:*
7: SP_Mhc:*
8: SP_Organelle:*
9: SP_Phage:*
10: SP_Plant:*
11: SP_Rodent:*
12: SP_Virus:*
13: SP_Vertebrate:*
14: SP_Unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1290	57.4	399	4	Q16529
2	1220.5	54.3	387	11	Q64194
3	1197.5	53.3	397	11	Q920M5
4	1134.5	50.5	395	11	Q9D798
5	1134.5	50.5	395	11	Q9D6X0
6	1134.5	50.5	395	11	Q9D606
7	1134.5	50.5	395	11	Q9D6P3
8	1133.5	50.4	395	11	Q9D6P8
9	1133.5	50.4	395	11	Q9D6P7
10	1131.5	50.4	395	11	Q9D6L9
11	1129.5	50.3	395	11	Q9D6T5
12	1128.5	50.2	395	11	Q9D6T7
13	1127.5	50.2	395	11	Q9D7C5
14	1126.5	50.1	395	11	Q9D766
15	1126.5	50.1	395	11	Q9D6L1
16	1125.5	50.1	395	11	Q9D6Q3
17	1124.5	50.0	395	11	Q9D760
18	1124.5	50.0	395	11	Q9D6S5
19	1122.5	50.0	395	11	Q9D796

20	1119.5	49.8	395	11	Q9D6N8	Q9d6n8 mus musculu
21	1116.5	49.7	395	11	Q9D6O2	Q9d6o2 mus musculu
22	1115.5	49.6	395	11	Q9D6M9	Q9d6m9 mus musculu
23	752.5	33.5	403	5	Q61866	Q61866 caenorhabd1
24	751	33.4	411	5	Q20449	Q20449 caenorhabd1
25	730	32.5	405	5	Q93789	Q93789 caenorhabd1
26	715	31.8	169	11	Q9D2L7	Q9d2l7 mus musculu
27	707.5	31.5	405	5	Q16956	Q16956 caenorhabd1
28	692	30.8	411	5	Q94252	Q94252 caenorhabd1
29	655.5	29.2	426	5	Q17766	Q17766 caenorhabd1
30	655.5	29.2	434	5	Q9VKR9	Q9vkr9 drosophila
31	639	28.4	456	5	Q9VKS5	Q9vks5 drosophila
32	636	28.3	398	5	Q9V796	Q9v796 drosophila
33	625.5	27.8	616	5	Q77107	Q77107 plodia inte
34	588.5	26.2	838	5	Q9V8K6	Q9vbk6 drosophila
35	574.5	25.6	559	5	Q17219	Q17219 bombyx mori
36	567.5	25.3	457	5	Q9VKR2	Q9vkr2 drosophila
37	559	24.9	399	5	Q9VPE9	Q9vpe9 drosophila
38	535	23.8	416	5	Q9VCO5	Q9vq5 drosophila
39	529.5	23.6	406	5	Q9VKT7	Q9vkt7 drosophila
40	527	23.5	504	5	Q94568	Q94568 galliera me
41	505	22.5	443	3	P78898	P78898 schizosacch
42	502	22.3	435	5	Q9VGA6	Q9vga6 drosophila
43	488	21.7	355	5	Q9VKR0	Q9vkr0 drosophila
44	485	21.6	356	5	Q9VKS9	Q9vks9 drosophila
45	478.5	21.3	401	5	Q9U276	Q9u276 caenorhabd1

ALIGNMENTS

RESULT 1
Q16529 PRELIMINARY; PTR; 399 AA.
AC Q16529
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE LYSOSOMAL ACID LIPASE PRECURSOR.
GN LAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid:9606;
RN [1]
RP TISSUE=LIVER;
RC SEQUENCE FROM N.A.
RA Du H., Gregory G.A.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
CC -i- SIMILARITY: TO ALPHA/BETA HYDROLASE FOLD.
DR EMBL; Z31690; CAA83495.1; -;
DR EMBL; U08464; AAB60328.1; -;
DR InterPro: IPR000073; Abhydrolase.
DR InterPro: IPR000379; Est_lip_thioest_actite.
DR InterPro: IPR000734; Lipase.
DR Pfam: PF00561; abhydrolase; 1.
DR PROSITE: PS00120; LIPASE_SER; 1.
KW Signal.
FT CHAIN 1 21 POTENTIAL.
FT VARIANT 22 399 LYSOSOMAL ACID LIPASE.
FT VARIANT 16 16 T->P (IN REF. 1).
FT VARIANT 23 23 R->G (IN REF. 1).
SQ SEQUENCE 399 AA; 45518 MW; AE051EE098A8152B CRC64;

Query Match 57.4%; Score 1290; DB 4; Length 399;
Best Local Similarity 61.0%; Pred. No. 2.8e-106;
Matches 242; Conservative 59; Mismatches 94; Indels 2;

QY 15 MEMLLIVAVYMFQRNNSVMPK--AVDPEAFMNISETIIOHGYPCEEYEATEGC
DB 1 MKMFLDLYVCLVMTLHSGSRGKLAVDPERMANVSEIISYGFSEETIVETEDC

QY	73	LSVNIPIPGGLVOPKKTGSRPVLLLOHGVGGASWNISLPGNLSGFLIADAGPDMWNGS	133
	1		
	61	LCINRIPIPGKRNKSHDKGPKPVFLQHGHLADSSMMVNTNLANSISGLFILLADGEDPMWNGS <th>1220</th>	1220
QY <th>133</th> <th>RGMAWSRKHKTLSDIDODEFMAFVSDEMARFDPALPVIINFILOKTGOEKIYYVGSQCTTMG</th> <th>192</th>	133	RGMAWSRKHKTLSDIDODEFMAFVSDEMARFDPALPVIINFILOKTGOEKIYYVGSQCTTMG	192
	121	RGNTWSRKHKTLVSODEDFMAFVSDEMAKYDLPASINFLINKTGOEDVYYVGHSGCTTIG	180
QY <th>193</th> <th>FIATSTMPWELAQTKIMFALAPITVTHAKASPGCKRFLLLPDMATKGLFGKKEPFIYQTRFL</th> <th>252</th>	193	FIATSTMPWELAQTKIMFALAPITVTHAKASPGCKRFLLLPDMATKGLFGKKEPFIYQTRFL	252
	181 <td>FIATSQDEPELAKRIKMFALGPVAVSFCTSPMAKLGRLPDHLTKDLFGKKEFLPOSATL</td> <th>240</th>	FIATSQDEPELAKRIKMFALGPVAVSFCTSPMAKLGRLPDHLTKDLFGKKEFLPOSATL	240
QY <th>253</th> <th>ROLYIYLCGGVILYILCOISINILILGGFTNTNMSRASVAAHTLAGTSYONILMSQAV</th> <th>312</th>	253	ROLYIYLCGGVILYILCOISINILILGGFTNTNMSRASVAAHTLAGTSYONILMSQAV	312
	241 <td>KMLGTHVCTHYILVELCGNLGFLLCGFERNRNLNSRDVYTHSPAGTSYONILMSQAV<th>300</th></td>	KMLGTHVCTHYILVELCGNLGFLLCGFERNRNLNSRDVYTHSPAGTSYONILMSQAV <th>300</th>	300
QY <th>313</th> <th>NSGELRAEDWMSSETKNLECKNOPTPVRYRVADMTVPTAMTGGQDMLSNBEDVYKMLLSEV</th> <th>372</th>	313	NSGELRAEDWMSSETKNLECKNOPTPVRYRVADMTVPTAMTGGQDMLSNBEDVYKMLLSEV	372
	301 <td>KEOKFQARDMSGSAKNVFHYNQSPRPVYNVADMLVPPAVMSGGDHLADYDVVILLTQI<th>360</th></td>	KEOKFQARDMSGSAKNVFHYNQSPRPVYNVADMLVPPAVMSGGDHLADYDVVILLTQI <th>360</th>	360
QY <th>373</th> <th>TNLIYKNIPEMAHVDFTWGLDAPHRMYNEIHLMOQ</th> <th>409</th>	373	TNLIYKNIPEMAHVDFTWGLDAPHRMYNEIHLMOQ	409
	361 <td>TNLVFHESIPMEHLDFTWGLDADAWRLNKLINLMRK<th>397</th></td>	TNLVFHESIPMEHLDFTWGLDADAWRLNKLINLMRK <th>397</th>	397
RESULT <th>2</th> <th></th> <th></th>	2		
Q64194 <th>Q64194</th> <th>PRELIMINARY;</th> <th>PRT; 397 AA.</th>	Q64194	PRELIMINARY;	PRT; 397 AA.
AC <th>Q64194;</th> <th></th> <th></th>	Q64194;		
DT <th>01-NOV-1996 (TReMBLrel_01, Created)</th> <th></th> <th></th>	01-NOV-1996 (TReMBLrel_01, Created)		
DT <th>01-NOV-1996 (TReMBLrel_01, Last sequence update)</th> <th></th> <th></th>	01-NOV-1996 (TReMBLrel_01, Last sequence update)		
DT <th>01-JUN-2001 (TReMBLrel_17, Last annotation update)</th> <th></th> <th></th>	01-JUN-2001 (TReMBLrel_17, Last annotation update)		
DE <th>LYSOSOMAL ACID LIPASE.</th> <th></th> <th></th>	LYSOSOMAL ACID LIPASE.		
OS <th>Rattus sp.</th> <th></th> <th></th>	Rattus sp.		
OC <th>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;</th> <th></th> <th></th>	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC <th>Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus</th> <th></th> <th></th>	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus		
OX <th>NCBI_TaxID=10118;</th> <th></th> <th></th>	NCBI_TaxID=10118;		
NP <th>[1]</th> <th></th> <th></th>	[1]		
RP <th>SEQUENCE FROM N.A.</th> <th></th> <th></th>	SEQUENCE FROM N.A.		
RX <th>MEDLINE-96129534; PubMed-8576647;</th> <th></th> <th></th>	MEDLINE-96129534; PubMed-8576647;		
RA <th>Nakagawa H., Matsubara S., Kuriyama M., Yoshidome H., Fujiyama J.,</th> <th></th> <th></th>	Nakagawa H., Matsubara S., Kuriyama M., Yoshidome H., Fujiyama J.,		
RA <th>Yoshida H., Osame M.;</th> <th></th> <th></th>	Yoshida H., Osame M.;		
RT <th>"Cloning of rat lysosomal acid lipase cDNA and identification of the</th> <th></th> <th></th>	"Cloning of rat lysosomal acid lipase cDNA and identification of the		
RL <th>mutation in the rat model of Wolman's disease.";</th> <th></th> <th></th>	mutation in the rat model of Wolman's disease.";		
CC <th>1-1. SIMILARITY: TO ALPHA/BETA HYDROLASE FOLD.</th> <th></th> <th></th>	1-1. SIMILARITY: TO ALPHA/BETA HYDROLASE FOLD.		
DR <th>EMBL; S81497; AAB36043.2; "</th> <th></th> <th></th>	EMBL; S81497; AAB36043.2; "		
DR <th>InterPro: IPR000073; Abhydrolase.</th> <th></th> <th></th>	InterPro: IPR000073; Abhydrolase.		
DR <th>InterPro: IPR000379; Est_lip_thioest-actsite.</th> <th></th> <th></th>	InterPro: IPR000379; Est_lip_thioest-actsite.		
DR <th>Pfam: PF00561; abhydrolase_1.</th> <th></th> <th></th>	Pfam: PF00561; abhydrolase_1.		
SO <th>SEQUENCE 397 AA; 45186 MW; 97A38595A0523947 CRC64;</th> <th></th> <th></th>	SEQUENCE 397 AA; 45186 MW; 97A38595A0523947 CRC64;		

[illegible]

Oy	272	I M L I G G E N T N H K S R A S V Y A A T L T A G T S V O U L I H M S O A N S E L P A F M G S E T K L E K	3391
Dd	258	I F F L I C G E N E K N L N S R D V Y T T C P A G T S Y O N N V H M T Q V Y K H K A D A F M G S S D K N Y F H	3177
Oy	332	C N O P T V R Y R D M T V P T A M T G O D M L S N E D Y K M L L E V T N L I Y H K N I P E M A H V P I W	3919
Dd	318	Y N O S T P P L Y S T K D M O L P A L M S G G K D M L A D I S I N I L T E I P L Y V I Y H K N I P E M D H L D F I W	3777
Oy	392	G L D A P H R M Y N E I I H L M O O	409
Dd	378	G L D A P W R L Y N E V S L M K	395

RESULT	3			
0920M5	ID	0920M5	PRELIMINARY;	PRT; 397 AA.
AC	0920M5			
DT	01-MAY-1999	(TREMBLrel. 10, Created)		
DT	01-MAY-1999	(TREMBLrel. 10, Last sequence update)		
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)		
DE	LYSOSOMAL ACID LIPASE PRECURSOR.			
GN	LIP1 OR LAL.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_Taxid=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57 BLACK/6X CBA; TISSUE=LIVER;			
RX	MEDLINE=96363957; PubMed=8725147;			
RA	Du H., Walte D.P., Grabowski G.A.;			
RT	"Tissue and cellular specific expression of murine lysosomal acid			
RL	lipase mRNA and protein."			
CC	J. Lipid Res. 37:937-949(1996).			
DR	-1- SIMILARITY: TO ALPHA/BETA HYDROLASE FOLD.			
DR	EMBL: Z31689; CAA83494.1; -			
DR	MGD: MGI:96789; Lip1.			
DR	InterPro: IPR000075; Abhydrolase.			
DR	InterPro: IPR000379; Est_lip_thloest-actsite.			
DR	InterPro: IPR000734; Lipase.			
DR	Pfam: PF00561; abhydrolase; 1.			
DR	PROSITE: PS00120; LIPASE_SER; 1.			
KW	Signal.			
FT	SIGNAL			
ET	CHAIN			
SO	SEQUENCE	397 AA;	45550. MW;	F886C39E1CCFA91F CRC64;

Query Match	53.3%	Score 1197.5	DB 11	Length 397
Similarity	55.6%	Pred. No. 4.5e-98		
Matches 217	Conservative 70	Mismatches 100	Indels 3	Gaps 1
Oy	23	VAYMFQRYNVSVMHPT---	KAVDPEAFNMISBIIHOGYPCBEYEYATEDGYYLSVNRIP	79
Db	6	LVFVETIGILLSRVYVTGYVSAYDPEVNNMVTIIMRWGYPGEHSVYIGDGYILSIHRIP		65
Oy	80	RGLVQPKRTGSRPVYLLHGLGVGASNNISLNNLSGFIILLADAGFDVWGMGSRGNAMSR		139
Db	66	RGMKHFQKGRPRVYLYHGLGLADSSNMVTINDNSLSGLFIILADRGFDVWGMGSRGNMTSL		125
Oy	140	KHKTIISIDODEFWASFYEDMARFDLPVAVINFLQKTGQEKIYYVGSQCTTYMGFLAFSTM		199
Db	126	KHKTSVSODEFWASFPEDMAKYDLPASINYLTKTGEOIYYVGHSSQCTTGIFLAFSQM		185
Oy	200	PELACKIMYFALAPAIAYVNAKSPGTRFLLPDMIMIGLPECKKEFIVQTRFLRDLVLYL		259
Db	186	PELAKIKIMFVLAPAVLISLNFASGPLDQGRPDLKLDMEFOKQFLPQSAIMLKMVLSHV		245
Oy	260	CGOVLDDIGCSINIMLLGCFNTNNMNSRASVYAHTLAGTSVQNILHMQSAVNSGELRA		319
Db	246	CTHYIMKEELCANVFLLDGGFNENKRNIMSGRVDVYTTTHCRPELLVQNMHLHGQVFFKRIKIDA		305
Oy	320	FDWGSERTNLEKCNQPTFVRYRVRDMTVPLAMTSGGDWLSNPEDVYKMLSEVTNLIYHK		379

[illegible]

Dd		101	FTSTNVLAKKIKRIFALPAPVATYKTKTESPFKISILIPKFLAVITGNKMFHNLDOPL	240
Oy		253	ROLVTYLGCQGVIVLDCSNIIMLLGGFNTNNMMSRASYAAHTAGTSVONILHNSOAV	312
Dd		241	GTEV---CSRELLDLCSSNALFIQCFGDKKNLVSRDLVGLGHNPAGTSTODLFHWAGLA	297
Oy		313	NSGEIARPDMSGETKNLECKNOPTFYRVRYRVOMTYPTAMTWGGOMLSNPEDVKMLISEY	372
Dd		298	KSGKIQAANNWSPDLNLMHYNKTPPYDYSAATPVIAVMNGSHILDADPODVAMLPLKL	357
Oy		373	TMLIYHKNPPEAHVDFTMGLDAPHRMVEIHLMQOE	410
Dd		358	PMLLYHKELLPYNHLDFIWAMPADPEVINEIVTMAMD	395
RESULT	5			
O9D6X0				
ID	O9D6X0	PRELIMINARY:	PRT:	395 AA.
AC	O9D6X0;			
DT	01-JUN-2001	(TREMBLrel. 17, Created)		
DI	01-JUN-2001	(TREMBLrel. 17, Last sequence update)		
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)		
DE	231005IB2IRIK	PROTEIN.		
GN	231005IB2IRIK.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.			
OX	NCBT_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=TONGUE;			
RX	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA	Aitawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,			
RA	Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Glass C., King B., Koichi H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,			
RA	Schirml L.M., Studli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA	Blake J., Boiffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustingich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,			
RA	Lions P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,			
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shihata J., Storch K.-F.,			
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,			
RA	Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsutkl S.,			
RA	Hayashizaki Y.;			
RL	"Functional annotation of a full-length mouse cDNA collection."			
CC	Nature 409:685-690(2001).			
CC	-1- SIMILARITY: TO ALPHA/BETA HYDROLASE FOLD.			
DR	EMBL; AK009875; BAB26556.1.			
DR	MGI; MG1:1914967; 2310051B2IRIK.			
DR	InterPro: IPRO00073; Abhydrolase.			
DR	InterPro: IPRO00379; Est_lip_thioest_actsite.			
DR	InterPro: IPRO00734; Lipase.			
DR	Pfam; PF00561; abhydrolase_1.			
DR	PROSITE; PS00120; LIPASE_SER; 1.			
SC	SEQUENCE 395 AA; 44665 MW; 40CA6B6759A8C5B CRC64;			
Query Match	50.5%; Score 1134.5; DB 11; Length 395;			
Best Local Similarity	53.5%; Pred. No. 1.8e-92;			
Matches 213; Conservative	70; Mismatches 108; Indels 7; Gaps 3.			
Oy	17	MMILLIVAYMFG-RMVNSVHMPTKAVDPFAFMPISEIIIOHOGPCPEEVATEEDGYI	75	
Dd	1	MMILLIVTVSLASFGGAHGILFKLGPKRKNEPANNAVSMITTYKGTIPSEIEIVYTEDEGITLAY	60	
Oy	76	NRIPGRVLQPPKKTGSRPVLLLOHGLVGASNMVISLPNNSLIGFILADAGFDVWMGNSRGN	135	

DR MGD:1914967; 2310051B21RIK.
 DR InterPro: IPR000073; Abhydrolase.
 DR InterPro: IPR000379; Est_lip_thioest_actsite.
 DR InterPro: IPR000734; Lipase.
 DR Pfam: PF00561; abhydrolase; 1.
 DR PROSITE: PS00120; LIPASE_SER; 1.
 SO SEQUENCE 395 AA; 44651 MW; E48A73CCCBFD359D CRC64;

Query Match 50.4%; Score 1134.5; DB 11; Length 395;
 Best Local Similarity 53.8%; Pred. No. 1.8e-92;
 Matches 214; Conservative 69; Mismatches 108; Indels 7; Gaps 3;

Dr 17 MMLILVAYMFQ-RNNSVHMPKAVDPEAFMNISETIQHOGYCEEXEYATEDEGYILSV 75
 1 MMLILVTSVLSAFGAGHLFGKLPKNEAMNYSOMITTYGYSSEYEYATEDEGYILSV 60
 Dr 76 NRIRGLVOPKKTGSRPVYLQHLGLVGASWMSINLPNNSIGFTLADAGDVMGNNSGN 135
 61 YRIPYGGKNSNIGKRPVAYLQHLGLASATWITNLPNNSLAFILADAGDVMGLNSGN 120
 Dr 136 AMSRHKRLSIDODEFAFSDYDEMARFDPVAVINFILOKTOGEKIYYGSGTGMFIA 195
 121 TWSRKNYYSPDSVEFMAFSDYDEMAKADLPATIDFVQKTOGEKIHYGSGTIGFIA 180
 Dr 196 FSTPDELAKTKMTFALAPFATVFKHNSPGTKFLLPDMITKGLFGKKEFL---YQREFL 252
 181 FSTPDELAKTKIRRYALAPFATVFKYTESPFKKISLIPKFLKLVFGNKMFPNHYLDLDFL 240
 Dr 253 RQLYIYLGQVILDOICSNIMLLGGFNTNMNNSRASVYAHTLAGTSYONILHWSQAV 312
 241 GTEV---CSRLLDLDCSNALFIFCGFDKKNLNSRFVYLGHNPDAGSTODLPHMAOLA 297
 Dr 313 NSGELRAFDSGSETRKLEKCNQPTPVRYRVDMVPTAMTGGODMISNPEDVKMLLSEV 372
 298 KSGRLQAVNMGSPLONMMLHYQKTPPYDVASAMTYPIAVNMGHDIADPDVAMLLPKL 357
 Dr 373 TNLVYHKNIPEMAVDFWGLDAPHRYMNETIHLMOOE 410
 358 PNLVYHKEILPYNHDFIWMADAPQEVYINELVTYMAED 395

RESULT 8.
 Dr 09CPE8 PRELIMINARY; PRT; 395 AA.
 AC 09CPE8;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE 2310051B21RIK PROTEIN.
 GN 2310051B21RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TONGUE;
 RX MEDLINE=21085660; Pubmed=11217851;
 RA Kawai J., Shinaagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 Schirml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohsaki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 CC -1- SIMILARITY: TO ALPHA/BETA HYDROLASE FOLD.
 DR EMBL: AK010116; BAB26711.1; -.
 DR EMBL: AK009300; BAB26201.1; -.
 DR MGD:1914967; 2310051B21RIK.
 DR InterPro: IPR000073; Abhydrolase.
 DR InterPro: IPR000379; Est_lip_thioest_actsite.
 DR InterPro: IPR000734; Lipase.
 DR Pfam: PF00561; abhydrolase; 1.
 DR PROSITE: PS00120; LIPASE_SER; 1.
 SO SEQUENCE 395 AA; 44579 MW; D229685EAE671D34 CRC64;

Query Match 50.4%; Score 1133.5; DB 11; Length 395;
 Best Local Similarity 53.5%; Pred. No. 2.2e-92;
 Matches 213; Conservative 70; Mismatches 108; Indels 7; Gaps 3;

Dr 17 MMLILVAYMFQ-RNNSVHMPKAVDPEAFMNISETIQHOGYCEEXEYATEDEGYILSV 75
 1 MMLILVTSVLSAFGAGHLFGKLPKNEAMNYSOMITTYGYSSEYEYATEDEGYILSV 60
 Dr 76 NRIRGLVOPKKTGSRPVYLQHLGLVGASWMSINLPNNSIGFTLADAGDVMGNNSGN 135
 61 YRIPYGGKNSNIGKRPVAYLQHLGLASATWITNLPNNSLAFILADAGDVMGLNSGN 120
 Dr 136 AMSRHKRLSIDODEFAFSDYDEMARFDPVAVINFILOKTOGEKIYYGSGTGMFIA 195
 121 TWSRKNYYSPDSVEFMAFSDYDEMAKADLPATIDFVQKTOGEKIHYGSGTIGFIA 180
 Dr 196 FSTPDELAKTKMTFALAPFATVFKHNSPGTKFLLPDMITKGLFGKKEFL---YQREFL 252
 181 FSTPDELAKTKIRRYALAPFATVFKYTESPFKKISLIPKFLKLVFGNKMFPNHYLDLDFL 240
 Dr 253 RQLYIYLGQVILDOICSNIMLLGGFNTNMNNSRASVYAHTLAGTSYONILHWSQAV 312
 241 GTEV---CSRLLDLDCSNALFIFCGFDKKNLNSRFVYLGHNPDAGSTODLPHMAOLA 297
 Dr 313 NSGELRAFDSGSETRKLEKCNQPTPVRYRVDMVPTAMTGGODMISNPEDVKMLLSEV 372
 298 KSGRLQAVNMGSPLONMMLHYQKTPPYDVASAMTYPIAVNMGHDIADPDVAMLLPKL 357
 Dr 373 TNLVYHKNIPEMAVDFWGLDAPHRYMNETIHLMOOE 410
 358 PNLVYHKEILPYNHDFIWMADAPQEVYINELVTYMAED 395

RESULT 9
 Dr 09CPE7 PRELIMINARY; PRT; 395 AA.
 AC 09CPE7;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE 2310051B21RIK PROTEIN.
 GN 2310051B21RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TONGUE;
 RX MEDLINE=21085660; Pubmed=11217851;
 RA Kawai J., Shinaagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirral L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Futuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guistincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Saeki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -1- SIMILARITY: TO ALPHA/BETA HYDROLASE FOLD.
 DR EMBL: AK010236; BAB26272.1; -
 DR EMBL: AK009413; BAB26272.1; -
 DR EMBL: AK009428; BAB26280.1; -
 DR EMBL: AK009459; BAB26300.1; -
 DR EMBL: AK009473; BAB26312.1; -
 DR EMBL: AK009474; BAB26313.1; -
 DR EMBL: AK009479; BAB26316.1; -
 DR EMBL: AK009523; BAB26338.1; -
 DR EMBL: AK009525; BAB26339.1; -
 DR EMBL: AK009546; BAB26332.1; -
 DR EMBL: AK009571; BAB26368.1; -
 DR EMBL: AK009573; BAB26370.1; -
 DR EMBL: AK009729; BAB26465.1; -
 DR EMBL: AK009773; BAB26466.1; -
 DR EMBL: AK010019; BAB26647.1; -
 DR EMBL: AK010035; BAB26656.1; -
 DR EMBL: AK010058; BAB26673.1; -
 DR EMBL: AK010061; BAB26675.1; -
 DR EMBL: AK010124; BAB26715.1; -
 DR EMBL: AK010125; BAB26716.1; -
 DR MGD: MGI:1914967; 2310051B21Rik.
 DR InterPro: IPR000073; Abhydrolase.
 DR InterPro: IPR000379; Est_lip_thioest_actsite.
 DR Pfam: PF00561; abhydrolase.1.
 DR PROSITE: PS00120; LIPASE_SER; 1.
 DR SEQUENCE 395 AA; 44637 MW; D3F96B5E6A71E34 CRC64;
 SO
 Query Match 50.4%; Score 1133.5; DB 11; Length 395;
 Best Local Similarity 53.5%; Pred. No. 2.2e-92;
 Matches 213; Conservative 70; Mismatches 108; Indels 7; Gaps 3;

QY 373 TNLVHKNPWAVDFIMGLDAPHRMYNEIHLMOOE 410
 DB 358 PNLVHKELPFLNHLDFIMADAPQEVNLETVMMAED 395
 RESULT 10
 ID Q9D6L9 PRELIMINARY; PRT; 395 AA.
 AC Q9D6L9;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE ADULT MALE TONGUE CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
 DE CLONE:2310076L13, FULL INSERT SEQUENCE.
 GN 2310051B21Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TONGUE;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Atakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirral L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Futuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guistincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Saeki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyshew-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -1- SIMILARITY: TO ALPHA/BETA HYDROLASE FOLD.
 DR EMBL: AK010203; BAB26766.1; -
 DR MGD: MGI:1914967; 2310051B21Rik.
 DR InterPro: IPR000073; Abhydrolase.
 DR InterPro: IPR000379; Est_lip_thioest_actsite.
 DR InterPro: IPR000734; Lipase.
 DR Pfam: PF00561; abhydrolase.1.
 DR PROSITE: PS00120; LIPASE_SER; 1.
 DR SEQUENCE 395 AA; 44671 MW; E8936162510A55C CRC64;
 SO
 Query Match 50.4%; Score 1131.5; DB 11; Length 395;
 Best Local Similarity 53.5%; Pred. No. 3.3e-92;
 Matches 213; Conservative 69; Mismatches 109; Indels 7; Gaps 3;

```

OY 253 ROLVYILCGOVLIDICSNIMLLGFEFTNNMMSRASYAAHTLAGTSVONILHMSQAV 312
DB 241 GREV---CSRELLDLICSNALFLFCGFDKKNLNSRFDVYLGNPAGTSTODLFHQAOLA 297
OY 313 NSGELRPFMGSEFTKLEKCNOPTPYRYVRDMPTPTAMTGGODMLSPEDVKMLSEY 372
DB 298 KSGKLOAYMWSPLQMLHYNOKSPPYDVSAVTPIAVWNGSHDILADPOVAMLLPRL 357

OY 373 TNLHYKKNPEWANHDFIWLGLDAPHRMNETHILMOOE 410
DB 358 PULLYHKEILPYNHDLFIWMDAPQEVNVEIYTMMAED 395

RESULT 11
OY 99D675 PRELIMINARY; PRT; 395 AA.
AC 09D675;
DB 01-JUN-2001 (TREMblrel. 17, Created)
DB 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DB 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE ADULT MALE TONGUE CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
DE CLONE:2310061A13, FULL INSERT SEQUENCE.
GN 2310051B21R1K
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosara H., Kondo S., Yamataka I.,
RA Alizawa K., Izawa M., Nishi K., Kiyosara H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schiml L.M., Staubl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
CC Nature 409:685-690(2001).
CC -I- SIMILARITY: TO ALPHA/BETA HYDROLASE FOLD.
DB EMBL: AK009990: BAB2629.1; -
DB MGD: MGI:1914967: 2310051B21R1K.
DB InterPro: IPR000073; Abhydrolase.
DB InterPro: IPR000379; Est_lip_thioest_actsite.
DB Pfam: PF00561; abhydrolase.1.
DB PROSITE: PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44623 MW; D3F96DB83161C3EF CRC64;

Query Match 50.3%; Score 1129.5; DB 11; Length 395;
Best Local Similarity 53.3%; Pred. No. 4,9e-92;
Matches 212; Conservative 71; Mismatches 108; Indels 7; Gaps 3;

OY 17 MWLLIVAYMFO-RNYSVHMPTRKAVDPAPFNISEIIOHGYPCREYVATEDGYILSV 75
DB 1 MWLLIVTAVLSAFGAGHGLFGKLGPKNPANNVNSQIMTYGVPSESEYEVVEDGYILGV 60
OY 76 NRIPIGLVQPKRTGSRPVYLLQHLVGLGASNNISLPPNSLGIITLADAFDVMGMSRGN 135
DB 1 NRIPIGLVQPKRTGSRPVYLLQHLVGLGASNNISLPPNSLGIITLADAFDVMGMSRGN 135

```

```

DB 61 YRIPYCKNSSENIQKRPVAYLQHLIASATNMTNLPPNSLFIILADAGDVWLNSRGN 120
OY 136 AMSRKHTLSIDODEFAPSYDEMARFDLPVAINFLQKGTGGEKIYYGSGITMGFLA 195
DB 121 TWSRKNVYVSPDSVEFMASPFDEMAKYDLPATIDFVQKGTGGEKIHYVHSGGTTIGFLA 180
OY 196 FSTMPLEAKIKMYFLATVATYKHAHKSPTGFTLLPDMWITGLPFCKEFL---VQTRFL 252
DB 181 FSTNPALKIRKIFVALAVATVRYKTESPPKISLIPKFLLYTGKMKFMPHNTLDQFL 240
OY 253 ROLVYILCGOVLIDICSNIMLLGFEFTNNMMSRASYAAHTLAGTSVONILHMSQAV 312
DB 241 GREV---CSRELLDLICSNALFLFCGFDKKNLNSRFDVYLGNPAGTSTODLFHQAOLA 297
OY 313 NSGELRPFMGSEFTKLEKCNOPTPYRYVRDMPTPTAMTGGODMLSPEDVKMLSEY 372
DB 298 KSGKLOAYMWSPLQMLHYNOKSPPYDVSAVTPIAVWNGSHDILADPOVAMLLPRL 357

OY 373 TNLHYKKNPEWANHDFIWLGLDAPHRMNETHILMOOE 410
DB 358 PULLYHKEILPYNHDLFIWMDAPQEVNVEIYTMMAED 395

RESULT 12
OY 99D67 PRELIMINARY; PRT; 395 AA.
AC 99D67;
DB 01-JUN-2001 (TREMblrel. 17, Created)
DB 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DB 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE 2310051B21R1K PROTEIN.
GN 2310051B21R1K
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Izawa M., Nishi K., Kiyosara H., Kondo S., Yamataka I.,
RA Alizawa K., Izawa M., Nishi K., Kiyosara H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schiml L.M., Staubl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
CC Nature 409:685-690(2001).
CC -I- SIMILARITY: TO ALPHA/BETA HYDROLASE FOLD.
DB EMBL: AK009537: BAB26346.1; -
DB MGD: MGI:1914967: 2310051B21R1K.
DB InterPro: IPR000073; Abhydrolase.
DB InterPro: IPR000379; Est_lip_thioest_actsite.
DB Pfam: PF00561; abhydrolase.1.
DB PROSITE: PS00120; LIPASE_SER; 1.
SQ SEQUENCE 395 AA; 44607 MW; E43317C2254FA8FB CRC64;

Query Match 50.2%; Score 1128.5; DB 11; Length 395;
Best Local Similarity 53.3%; Pred. No. 6.1e-92;

```

Matches 212; Conservative 70; Mismatches 109; Indels 7; Gaps 3;

QY 17 MMLILVAFMQ-RNVNSVHMPKAVDPAPFNNISIIIOHGPCPEYEVATEDGIIISV 75
 |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : :
 Db 1 MMLILVTSVLSAFGAGHGLFGKLGPKNPANNAVSOIMTYWGVPSEYEYVEDGIIISV 60

QY 76 NRIPRGLOPKKTGSRPVVLLQHGLVGASNMISNLPNNSLGFIILADAGFDVMMGNSRGN 135
 |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : :
 Db 61 YRIPYKKNSENIGKRPVAVIQLHGLIASAKNNITNLPNNSLAFIILADAGFDVMMGNSRGN 120

QY 136 AMSRKHKTLSIDODEFAESYDEMARFDPAYINFILOKTGGEKIYVYSGGTGFI 195
 |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : :
 Db 121 TMSRKNVYSPDSVEFWAFSPDEMAKYDLPATIDFVQKTGGEKIHYVGHSGGTGFI 180

QY 196 FSTMPLEAKIKMYFALPIATVYKAAKSPGTFLPDMKIKLFGSKKEFL--YOTRLL 252
 |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : :
 Db 181 FSTNPALAKKIRFYALAVATVYKTESPFKISILPKRLKVIYFNKKNFMHNYIDQPL 240

QY 253 ROLVYILGQVILDOICSNIMLLGGFNTNNMNSRASYAAHTLAGTSVONILHMSQAV 312
 |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : :
 Db 241 GTEV--CSRELLDLCSNLFIFCGFDKNNLVSRFDVYLGHNPAGTSTODLFHWAQLA 297

QY 313 NSGELRPFMGSETKNLEKNOPTPYRYRVDMTVPFAMTGGQDMLSPEDVKMLLSEY 372
 |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : :
 Db 298 KSGKLOAYNMWSPLOMMLHYNOKTPPYDVSAATVPVIAVWNGHDLADPODVAMLLPRL 357

QY 373 TNLIIHKNIPEVAHVDFTGLDAPHRVNEIHLMOOE 410
 |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : :
 Db 358 PNLVYHKEILPYNHDFIWMADAPDEVYNEIYMAED 395

RESULT 13
 Q9D7C5 PRELIMINARY; PRT: 395 AA.

AC Q9D7C5: 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE 2310051B21RIK PROTEIN.
 GN 2310051B21RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TONGUE;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Aizawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana S.,
 Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 Kuehl P., Lewis S., Matsuo Y., Nikaio I., Pesole G., Quackenbush J.,
 Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 Norioka P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
 Wyszyn-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
 Hayashizaki Y.;
 RA *Functional annotation of a full-length mouse cDNA collection.*;
 RL Nature 409:685-690(2001).
 CC -i- SIMILARITY: TO ALPHA/BETA HYDROLASE FOLD.
 DR EMBL: AK009359; BAB26240.1;
 DR MGD: MGI:1914967; 2310051B21RIK.
 DR InterPro: IPR000073; Abhydrolase.
 DR InterPro: IPR000379; Est_lip_thioest_actsite.

DR InterPro: IPR000734; Lipase.
 DR Pfam: PF00561; abhydrolase.1.
 DR PROSITE: PS00120; LIPASE_SER; 1.
 SQ SEQUENCE 395 AA; 44748 MW; 3ED43992458DE058 CRC64;

Query Match 50.2%; Score 1127.5; DB 11; Length 395;
 Best Local Similarity 53.3%; Pred. No. 7.4e-92;
 Matches 212; Conservative 69; Mismatches 110; Indels 7; Gaps 3;

QY 17 MMLILVAFMQ-RNVNSVHMPKAVDPAPFNNISIIIOHGPCPEYEVATEDGIIISV 75
 |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : :
 Db 1 MMLILVTSVLSAFGAGHGLFGKLGPKNPANNAVSOIMTYWGVPSEYEYVEDGIIISV 60

QY 76 NRIPRGLOPKKTGSRPVVLLQHGLVGASNMISNLPNNSLGFIILADAGFDVMMGNSRGN 135
 |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : :
 Db 61 YRIPYKKNSENIGKRPVAVIQLHGLIASAKNNITNLPNNSLAFIILADAGFDVMMGNSRGN 120

QY 136 AMSRKHKTLSIDODEFAESYDEMARFDPAYINFILOKTGGEKIYVYSGGTGFI 195
 |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : :
 Db 121 TMSRKNVYSPDSVEFWAFSPDEMAKYDLPATIDFVQKTGGEKIHYVGHSGGTGFI 180

QY 196 FSTMPLEAKIKMYFALPIATVYKAAKSPGTFLPDMKIKLFGSKKEFL--YOTRLL 252
 |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : :
 Db 181 FSTNPALAKKIRFYALAVATVYKTESPFKISILPKRLKVIYFNKKNFMHNYIDQPL 240

QY 253 ROLVYILGQVILDOICSNIMLLGGFNTNNMNSRASYAAHTLAGTSVONILHMSQAV 312
 |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : :
 Db 241 GTEV--CSRELLDLCSNLFIFCGFDKNNLVSRFDVYLGHNPAGTSTODLFHWAQLA 297

QY 313 NSGELRPFMGSETKNLEKNOPTPYRYRVDMTVPFAMTGGQDMLSPEDVKMLLSEY 372
 |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : :
 Db 298 KSGKLOAYNMWSPLOMMLHYNOKTPPYDVSAATVPVIAVWNGHDLADPODVAMLLPRL 357

QY 373 TNLIIHKNIPEVAHVDFTGLDAPHRVNEIHLMOOE 410
 |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : : |||||: : : : :
 Db 358 PNLVYHKEILPYNHDFIWMADAPDEVYNEIYMAED 395

RESULT 14
 Q9D766 PRELIMINARY; PRT: 395 AA.

AC Q9D766: 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE 2310051B21RIK PROTEIN.
 GN 2310051B21RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TONGUE;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Aizawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana S.,
 Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 Kuehl P., Lewis S., Matsuo Y., Nikaio I., Pesole G., Quackenbush J.,
 Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 Norioka P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
 Wyszyn-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
 Hayashizaki Y.;
 RA *Functional annotation of a full-length mouse cDNA collection.*;
 RL Nature 409:685-690(2001).
 CC -i- SIMILARITY: TO ALPHA/BETA HYDROLASE FOLD.
 DR EMBL: AK009359; BAB26240.1;
 DR MGD: MGI:1914967; 2310051B21RIK.
 DR InterPro: IPR000073; Abhydrolase.
 DR InterPro: IPR000379; Est_lip_thioest_actsite.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2002, 20:24:53 ; Search time 1858.36 seconds
(without alignments)

11265.267 Million cell updates/sec

Title: US-09-578-063-46

Perfect score: 1269
Sequence: 1 atgttggaacctgtctcag.....gacggtgtgagcgccgattg 1269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 segs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_com: *
20: em_or: *
21: em_ov: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_sy: *
28: em_un: *
29: em_vl: *
30: em_htgo_hum: *
31: em_htgo_inv: *
32: em_htgo_rod: *
33: em_htg_hum: *
34: em_htg_inv: *
35: em_htg_rod: *
36: em_htg_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	527.8	41.6	2481	9 HSLYACLY	Z31690 H. sapiens (
2	527.8	41.6	2481	9 HSLYACLY	U08464 Human Lysos
3	527.8	41.6	2481	9 HSLYACLY	X76488 H. sapiens m
4	526.2	41.5	2493	9 HUMILPCHL	M74775 Human Lysos
5	524.6	41.3	2586	9 BC012287	BC012287 Homo sapi
6	489.2	38.6	1378	6 A26689	A26689 Precursor o
7	489.2	38.6	1378	6 A26689	A26689 Precursor o
8	470	37.0	1365	9 HSGLR	X05997 Human mRNA
9	470	37.0	1367	6 A01046	A01046 H. sapiens m
10	467.6	37.0	1367	6 A12714	A12714 Pregastric
11	467.6	36.8	1336	6 A01157	A01157 R. norvegicu
12	467.6	36.8	1335	10 RNLP	X02309 Rat mRNA fo
13	461	36.3	2358	10 MMLYACLY	Z31689 M. musculus
14	459.8	36.2	3144	10 S81497	S81497 Lysosomal a
15	458	36.1	1417	4 BOVME	I26319 Bos taurus
16	444.2	35.0	1137	6 AR039022	AR039022 Sequence
17	444.2	35.0	1137	6 AR092633	AR092633 Sequence
18	444.2	35.0	1140	6 AR038023	AR038023 Sequence
19	444.2	35.0	1146	6 AR092634	AR092634 Sequence
20	444.2	35.0	1146	6 AR039024	AR039024 Sequence
21	444.2	35.0	1146	6 AR092635	AR092635 Sequence
22	444.2	35.0	1198	6 A57760	A57760 Sequence 5
23	444.2	35.0	1528	6 A57756	A57756 Sequence 1
24	444.2	35.0	1528	6 AR038021	AR038021 Sequence
25	444.2	35.0	1528	6 AR092632	AR092632 Sequence
26	444.2	35.0	1651	4 CFLIPASE	Y13899 Canis fam11
27	441	34.8	1137	6 A39301	A39301 Sequence 2
28	441	34.8	1140	6 A39303	A39303 Sequence 4
29	441	34.8	1146	6 A39305	A39305 Sequence 6
30	441	34.8	1528	6 A39300	A39300 Sequence 1
31	388.8	30.6	1048	6 A57758	A57758 Sequence 3
32	286	22.5	699	6 AX068249	AX068249 Sequence
33	286	22.5	1308	6 AX068259	AX068259 Sequence
34	269	21.2	43456	2 AL353113	AL353113 Human DNA
35	269	21.2	76653	2 AC015506	AC015506 Homo sapi
36	269	21.2	76653	2 AC015506	AC015506 Homo sapi
37	204.4	16.1	148720	2 AC019353	AC019353 Homo sapi
38	204.4	16.1	172834	2 AC011277	AC011277 Homo sapi
39	151.2	11.9	76653	2 AC015506	AC015506 Homo sapi
40	147	11.6	161826	9 AL358532	AL358532 Human DNA
41	140.6	11.1	401	9 HSLIPCHL04	U04288 Human Lysos
42	140.6	11.1	1851	9 HSLIP4	X75491 H. sapiens L
43	140.6	11.1	169404	2 AL353751	AL353751 Homo sapi
44	140.6	11.1	188147	2 AL353753	AL353753 Homo sapi
45	127.6	10.1	40114	3 CELZK6	AF067942 Caenorhab

ALIGNMENTS

RESULT	1	
LOCUS	HSLYACLY	2481 bp mRNA
DEFINITION	H. sapiens (HepG2) LAL mRNA for lysosomal acid lipase.	17-FEB-1997
ACCESSION	Z31690	
VERSION	Z31690.1	GI:506430
KEYWORDS	LAL; lipase; lysosomal acid lipase.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE	Structural Conservation of Putative Functional Motifs between Mouse and Human Lysosomal Acid Lipase	
JOURNAL	Unpublished	
REFERENCE	Du, H. and Gregory, G.A.	
AUTHORS	2 (bases 1 to 2481)	
TITLE	Du, H.	
JOURNAL	Submitted (05-APR-1994) Hong Du, Division of Human Genetics,	

Children's Hospital Medical Center, 3333 Bernet Street, Cincinnati,
OH, 45229-3039, USA

FEATURES

source

Location/Qualifiers

1. .2481

/organism="Homo sapiens"

/strain="HepG2"

/db_xref="taxon:9606"

/clone="clone hLAL"

/tissue_type="liver"

/cell_type="hepatoma"

/clone_lib="human liver HepG2 cDNA library/lambda Zap"

34. .96

/gene="LAL"

34. .1233

/gene="LAL"

34. .1233

/gene="LAL"

/citation=[1]

/codon_start=1

/product="lysosomal acid lipase"

/protein_id="CA83495.1"

/db_xref="GI:506431"

/db_xref="SPTREMBL:O16529"

/translation="MKKRFGLVYCVIYMLTHSGSRKLTAVDPENMVSEISYV
GPFSEELVEEDGYILCLNRIPIHGRKNSDKRPVYLOHGLADSSKWTNLAN
SLGFIILADAGFDVWNGNSRNTWSRKHTLSVSDDEFWATSYDEMAVYDI.PASINFL
NKTGOEYVYVGSOGTTIGFIASFQ.IPELAKRKMEFAPVSAVFCSPAKLGR
LPDLIKDLGDEKLEFLPOSALFKLGTHTVHLKELCNLCFLCGFERNLNMR
VDVYTHSPAGTSVONMLHWSNAVKFQKPAFDMGSSAKNVEFHNSYPPYVNDKL
VPIAVNSGGHMDLADYDVNILLTQITNLYFHESIPEMHLDPITMGIDAPWRLNKII
NLMRKIV"

79

/gene="LAL"

/citation=[1]

/replace="C"

97. .1230

/gene="LAL"

/product="lysosomal acid lipase"

100

/gene="LAL"

/citation=[1]

/replace="G"

2459. .2464

polyA_signal

polyA_site

2480

BASE COUNT

692 a 477 c 513 g 799 t

ORIGIN

Query Match 41.6%; Score 527.8; DB 9; Length 2481;

Best Local Similarity 65.4%; Pred. No. 7.3e-147;

Matches 791; Conservative 0; Mismatches 412; Indels 6; Gaps 1;

OY 39 cagaatgaatggtgcttcgattcgtggtggtatattgttcagagaatgtaattc 98
DB 30 CAGATGGAATGCGGTCTTGCGGTGTGCTGTTGTTCTCTGAGACCCCTCATTC 89
OY 99 ag-----tacatgccaactaaagcgtgtgaccagaagcatcagaatattcga 152
DB 90 TGAGGGGTCTAGAGGGAACACTGCAGCTGTGATCTGAACAACTGAATGTAGTGA 149
OY 153 aatcacaacatcaaggctacccctgtgagaatataatgacatgaagatgagta 212
DB 150 AATTAATCTTACTGGGATTCCTCCTAGAGAAATACCTAGTTGAGACAGAGATGATA 209
OY 213 tatcttctctgttaacagattccctcagagctagtgcacaactaagaagacggttcag 272
DB 210 TAATTTGTGCTTAAACCAATTCCTCATGAGAGAAACCATTCCTGACAAAGTCCCA 269
OY 273 gccctggtgttactgcagcatggtcgtgaggtgtcagcaactgatttcacact 332
DB 270 ACCAATTGTCTTCTGACACATGCGCTTGTGTCAGATCTTACTAATCTGGGTACAAACT 329
OY 333 gccacaacatagcctggtcttcattctcgtcagatgctgttttgacgtgtgagga 392

DB 330 TGCCACACAGACGCTGGGCTTCATCTTGGTGAATGCTGTTTACCTGTGGATGGCA 389
OY 393 cagcaggaggaacgcctggtctcgaacacacagacacatccatcgaagatggtt 452
DB 390 CAGCAGAGGAATATCTGGGTCTCGGAATTAACACACTCTCATTTCTGAGTAAT 449
OY 453 ctggtcttcaagtaatgagatgctgaggttggtaccctccgcaatgataactat 512
DB 450 CTGGGCTTTCAGTATGATGAGATGCAAAATATGACCTTACACCTTCACTTAT 509
OY 513 ttgcagaaacagcggcagaagaaatcattatgctggtcattccagagccacacat 572
DB 510 TCTGAATTAACCTGGCCAGACAGATGTATTATGTGGGTATTTCTCAAGCACACTAT 569
OY 573 ggccttattcattccacatgacagatgctgctcagaacaaatcgaatgtttgc 632
DB 570 AGGTTTATAGCATTTTACAGATCCCTGACCTGGCTTAAGATTAATGTTTTC 629
OY 633 tttagcaccatagccactgttaagcatgcaaaagcccgagacaaattttgtgct 692
DB 630 CCTGGGCTCTGTGGCTTCGCGCTTCTGACTAGCCCTATGCGCAATATGACAGATT 689
OY 693 gccagatatgatatgaggaattgtttgcaaaaagaatttctgtatcagaccagatt 752
DB 690 ACCGATCATCTCATTTAAGACTTATTTGAGACAAAGAAATTTCTCCCGAGATGCTT 749
OY 753 tctcagaacactgttattacccttctgycgaagtgattcctgatcagattgataa 812
DB 750 TTTGAAGGCTGGGTACCCACGTTTGACACATGCTCATGTAAGAGGCTCTGGAAA 809
OY 813 tatcatgttactctggtgtgattccaacacacacataatgacatgagcgaagatgt 872
DB 810 TCTCTGTTTCTCTGCTGTGATTTAAGAGAAATTAATATGCTTAAGAGGATGT 869
OY 873 atagctgcccacactctgttgtaacactcgttcgaatattctacacttgagccagc 932
DB 870 AATATCAACACATTTCTCTGTAACCTCTGTGAAAACATGTTACCTGAGCCAGGC 929
OY 933 agtgaattcgtgtaactccggcattgacttggtggtgagatgagacaaactcgtgaaa 992
DB 930 TGTATTAATCCAAAAGTTTCAAGCTTGTGATCGGGAAGCATGTCAGAAATTTTCA 989
OY 993 atgcaatcagcacaactcctgtaagtaagtaagtaagatgacgtgacccaagcaat 1052
DB 990 TTACACACAGATTATCTCCACATACATGATGTAAGACATGCTTGTGCCAGTGCAGT 1049
OY 1053 gtgagacagaggtcagagctgcttcaatccagaagacgtgaaatgctgctctga 1112
DB 1050 CTGGAGCGGGGTGACGACTGGCTTGAGATGCTTACAGACGTCATATCTTACTACTCA 1109
OY 1113 gttgacacactcattccatlaagaatattcctgaatggtcgcgtggaattcattctg 1172
DB 1110 GATCAACCACTTGTGTCTTCAAGAGATTTCCGAATGGAGATCTTGACTTCATTG 1169
OY 1173 ggtgttgatctctcctcagctatgtaacatgaaatcattcattcgtgacgagga 1232
DB 1170 GGGCCTGATCCCTCTGAGAGCTTTAATAAATTTATATCAATAGAGAAATATCA 1229
OY 1233 gaccacact 1241
DB 1230 GTGAAGCT 1238

RESULT 2
LOCUS HSU08464 2481 bp mRNA . PRI 23-JUN-1994
DEFINITION Human lysosomal acid lipase mRNA, complete cds.
ACCESSION U08464
VERSION U08464.1 GI:505052
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL
1	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	(bases 1 to 2481)	
2	Du, H. and Gregory, G. A.	Structural Conservation of Putative Functional Motifs between Mouse and Human Lysosomal Acid Lipase	Unpublished (bases 1 to 2481)

TITLE Direct Submission
JOURNAL Submitted (08-APR-1994) Hong Du, Division of Human Genetics,
Children's Hospital Medical Center, 3333 Burnet Street, Cincinnati,
OH 45229-3039, USA

FEATURES	Location/Qualifiers
source	1. .2481

5'UTR
sig_peptide
CDS

```

/codon_start=1
/product="lysosomal acid lipase"
/protein_id="AAB60328.1"
/db_xref="GI:505053"

```

translation "KKMRKRLCPLGCVLWMTLHSEGRKGLTAVDPENNNWVLSIYW
GPPSEVLETEDGTYITLCNLRIPHGKRNNSDKDQKPLFPHOHLDDNNWVLSIYW
SLGLFLLADGAVGPMGNTSGNNGSRKRLKSLVSDQDEPA SYEMAKYDIPASTNPL
NKTGOEYVYVHSGOATGTGFLPAISOPELAKRIFALGPAVSAFCSPAKRGR
LPDHLIKDLGFEKTLPOAGFLKCTHCTHATYIKELCQDIFLFCGFEYNNYNSK
VDVYTHHSAGSYONELMSOAVKRCFOAFPMWGSACKNYFAYNISOPTYNNKDL
VITPAWMSGHMLADVYVNIILTLQTINLPHESIPEMWHLDFIWDGADWRLNKLII
NIMRKYQ"

	variation	79	/replace="c"	
mat_peptide		97	.1230	
				/product="lysosomal acid lipase"
variation		100		
				/replace="g"
3'UTR		1238	.2480	
polyA_signal		2459	.2464	
polyA_site		2480		
BASE COUNT	692 a	477 c	513 g	799 t
ORIGIN				

Query Match	41.6%	Score 527.8;	DB 9;	Length 2481;
Best Local Similarity	65.4%	Pred. No. 7.3e-147;		
Matches 791; Conservative	0;	Mismatches 412;	Indels 6;	Gaps 1;

OY	39	cagaatggaattcgtggcttcgcgatcccgagcgcatatcttccaaggaaatgtaac	98
Db	30	CAGATGAAATGGCGTTCTTGSGGTGGTGGTCTGTTGGTTCCTCGAACCCGCATT	89
OY	99	ag-----tacatatgccaaactaaagcgtgynaccgaacalcataatagtga	152
Db	90	TGAGGGGTCTAGAGGGAACAGACGCTGTGGACTCTGAAACAACATGATGTGACTGA	149
OY	153	aatcatcaaacatcaaggctatccctcgttgaggaaatagaagtcgcacctaagaatggtta	212
Db	150	AATATATCTTACTGGGGATTCCTTAATGAGGAATACCCTGTTGAGACAAAGAATGGATA	209
OY	213	tatcccttcgttaacaaggatctccctcgaagccctagtcycaacctaaagaagaagtttcag	272
Db	210	TATCTGTGGCTTAACCGAATTCCTCAATGGGAGGANAACACATTCTCACAAAGTCCAA	265
OY	273	gcctcgtgctgtactgcagcaatgagcctagtcttgaaagtgctagcaactggaattccaact	332
Db	270	ACCAGTTGTCTCCTCGCAACATGGCTTGCAGGCGAATTCAGTAACAGGGGTCAACAAC	329

QY	333	gccacaataagcctgggtcttcattcttgcgagaatgcgtgttttgaaagtgtgaatggaa	392
Db	330	TTGCAACAGCAGCCCTGGGCTTCATCTTCCTGTAGTGTGGTTTGGAGCTGTGATGGGAA	389
QY	393	cagcaagggaagacgcctgggtctcgaaaacacagacactctccatctagaccgaatggat	452
Db	390	CAGCAGAGGAATACCTGGTGTCTCGAANACATAAGACACTCTCAGTTTCTCAGAGTAAT	449
QY	453	ctgggtcttcagttatgatagatgctgaagtgttagacttccttcgacgtatataacttat	512
Db	450	CTGGGCTTTCAGTTATGATGATGATGAGTGGCAAAATATGACATACAGCTTCCATTAACTCAT	509
QY	513	tttcagaaaaacggcgccaggaaaaagatcttatgttcggtatltcaagggaccaccat	572
Db	510	TCTGAATTAACCTGGCCAAACAAAGTATATATGTGGGTCAATTCTCAAGGCACCATAT	569
QY	573	gggcttatcttcattctccacacatgcgagagcctgggtctcgaagaatccaagtattcttc	632
Db	570	AGGTTTATACATTTTTCACAGATCCCTGAGCTGGCTAAAGGATTAAATTTTTTTGC	629
QY	633	tttaagccccaatgaaccactgtlaagcatgacgaaaaagccccggagaccaaatttttgtc	692
Db	630	CTGGGTCTCTGGCTTCGGTCGCGCTTCTGTACTAGCCCTATGCGCAATATGAGCAGT	689
QY	693	gccagatatgatcatgaaggatgtttgtgcaaaaaagaattcttgtatcagaccaagat	752
Db	690	AACGATCATCTCATTTAAGGACTTATTTGGAGCAAAAGATTTTCTCCACAGTGCCTT	749
QY	753	ttcgaagcaactgtttatcactttgttgccagggatcttcagacagattgaa	812
Db	750	TTTGAAGTGGCTGGGTACCCACGTTTGCACTCATGTCACTGAGGAGACTTGTGGAA	809
QY	813	tatcatgtactcttgggtgattatcaacacaaacaaatagacatgagccgacaagt	872
Db	810	TCTCGTTTCTTCTGTGTGATTTAATAGAGAAATTTAAATATGTCTAGTGGATGT	869
QY	873	atatgtctgcacacactcttgcctggaacatctgtgcaaaaatlttaactctgagccagc	932
Db	870	ATAATACAAACATTTCTCTCTGTGAACTTCTGTGCAAAACATGTTTACCTGGAGCCAGCG	929
QY	933	agtaattctcgtgaactccgggcatttgcctggggagttgagccaaaatctcga	992
Db	930	TGTTAATTCACAAAGTTTAAAGCTTTGACTGTGGAGACGACTGCCAAGATTAATTTC	989
QY	993	atgcaatcagccaactctcttgaagtacagagtcagagatatgaagttccctacagcaat	1052
Db	990	TTAACAACAGATTAATCTCCACATACATGATGAAGGACATGCTGTGGCCACTGCAGT	1049
QY	1053	gtgcaacaggaagtcaggaactcgtcttcaaatccagaagaagtgaaatgctcctcga	1112
Db	1050	CTGAGAGGGGGGTACACACTGCGCTTCAGATGTCTCGACGTCATATCTTACTACATCA	1109
QY	1113	gtgtaaccaactctatcacataagaatattcctgaatgggttcacggtattcatctg	1172
Db	1110	GATACCAACTGTGTGTTCCATAGAGCATTCGCCGATGGAGCATTTTGACTTCAATTG	1169
QY	1173	gggtttgattctctccacccgtatgtacatgaatcaatccatctcgtatgcagcagga	1232
Db	1170	GGGCGCTGGATGCCCTTGGAGGCTTTAATATAAATTAATTAATTAATGAGGAATAATCA	1229
QY	1233	gaccacact 1241	
Db	1230	GTGAAGCT 1238	

RESULT	3				
LOCUS	HSLLAL				
DEFINITION	HSLLAL	2626 bp	mRNA		
ACCESSION	H.sapiens mRNA for Lysosomal acid lipase.			PRI	25-FEB-1994
VERSION	X76488				
KEYWORDS	X76488.1 GI:434305				
	lysosomal acid lipase.				

SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2626)
 AUTHORS Ameis, D., Merkel, M., Eckerskorn, C. and Greten, H.
 TITLE Purification, characterization and molecular cloning of human hepatic lysosomal acid lipase
 JOURNAL Eur. J. Biochem. 219 (3), 905-914 (1994)
 MEDLINE 94155897
 REFERENCE 2 (bases 1 to 2626)
 AUTHORS Ameis, D.
 TITLE Direct Submission
 JOURNAL Submitted (29-NOV-1993) D. Ameis, Medical Department, University Hospital Eppendorf, Martinistrasse 52, 20246 Hamburg, FRG
 FEATURES
 source
 1. 2626
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="10"
 /tissue_type="liver"
 146..1345
 /gene="LAL/LIPA"
 146..1345
 /gene="LAL/LIPA"
 /EC_number="3.1.1.13"
 /note="sterol esterase"
 /codon_start=1
 /evidence="experimental"
 /product="lysosomal acid lipase"
 /protein_id="CA54026.1"
 /db_xref="GI:434306"
 /db_xref="SWISS-PROT:P38571"
 /translation="MKMRLGLVCLVLTMLHSESGKLTAVDEPTNNVSEIISYM
 GPESEYLVETEDGYILCLRIIPHGKNSDKPKPVVEFLQHLADSSNWTNLNS
 SIGFLIADGPFVWNGNSRNTMSRKHTLVSODEPMAFSYDEMAKXDLPAISNIFL
 NKTCGOQVYVYHGSQCTTGFIATFSOIPELAKRIKFPALGPASVAFCTSPAKIKR
 LPDHLTKDLFGDKERLPQSAFLKWLSTHCTHYILALEGNLCLFLCGFNERNLNRSR
 VDVYTHSPAGTSVQNMHLMSQAVKFOKQAFDWGSSANRYHNSYPTPVNKKML
 VPTAVWSGGHMDLADYVDVNIILTOITNLVFEHSIPWEHLDFIMGLDAPWRLNKKII
 NLMKRYQ"
 146..226
 /gene="LAL/LIPA"
 /evidence="experimental"
 227..373
 /gene="LAL/LIPA"
 /note="propeptide"
 /evidence="experimental"
 374..1342
 /gene="LAL/LIPA"
 /evidence="experimental"
 /product="lysosomal acid lipase"
 mat_peptide
 BASE COUNT 741 a 527 c 547 g 811 t
 ORIGIN
 Query Match 41.6%; Score 527.8; DB 9; Length 2626;
 Best Local Similarity 65.4%; Pred. No. 7.4e-147;
 Matches 791; Conservative 0; Mismatches 412; Indels 6; Gaps 1;
 Oy 39 caaataagaaatgtgcttgaatctgtgtggcgatataatccagaagaatgaaatc 98
 Db 142 CAGAAATGAATAATGCGGCTTCTGGGCTGTGGCTGTTGTTCTCTGACCCCTGCATTC 201
 Oy 99 ag-----tacatatgcaactaaagctgtgagcccaagaacatcattgataatag 152
 Db 202 TGAAGGCTGTGAGGAACTGACAGCTGTGATCTCGAACAACAATGATGATGATGA 261
 Oy 153 aatcatcaacaatcaagctatccctgtgaggaataatgaaatcgcaactgaagatg 212
 Db 262 AATTATCTTACTGTGGGATTCCTACTGAGAAATACCTGTTGAGACAGAAAGATGATA 321
 Oy 213 tatccttctgttaaacaggtctcctgagccctagtcgaacctaaagaagcaggttcag 272

Db 322 TATTCTGTCCTTAACCGAATTCCTCATGGAGAGAAACAACTCTGACAAAGTCCAA 381
 Oy 273 gacctgtgttctactgcaagcatgacctagtgtgaagtgctcaagaactgattccaact 332
 Db 382 ACCAGTGTCTCTCTGCAACATGCTGCTGCGAGATTCAGTAACAGGCTCCAAACCT 441
 Oy 333 gcccaacaatagctgtgacctcattctgtgcagatgtcgtgttgaagctgtgtgag 392
 Db 442 TGCCAACAGCAGCCTGGGCTTCATTCCTGCTGATGCTGTTGACCTGTGATGGCA 501
 Oy 393 cagcaggggaaacgcctgtctgaaacaagaacacatccatagaccacaagaatgatt 452
 Db 502 CACGACGAGAAATACCTGCTGCTGGAACATTAAGACACTCTCAGTTCTCAGATGAA 561
 Oy 453 ctggcttctcattatgatatgatatgctgtagtttgacctctgcagtgataacttatt 512
 Db 562 CTGGGCTTTCAGTTATGATGATGAGCAAAATATGACCTACAGCTTCCTTAACCTCAT 621
 Oy 513 ttgcagaaaaagggccagaagaaagatctattatgtcgtctattcaagggcaccacat 572
 Db 622 TCTGAATAAACCTGGCCAAAGCAAGTATTTATGTGGTCATTCCTAAAGCACACTAT 681
 Oy 573 gggcttattgcatcttccaccatgcccagagctggctcagaagaaatcaaatgtatttgc 632
 Db 682 AGGTTTATAGCATTTTTCACAGATCCCTGAGCTGCTTAAGATTAATGTTTTTTCG 741
 Oy 633 tttagcaccatagaccactgtttaagcatgcaaaagcccgaggccaacttttctgtc 692
 Db 742 CCGGGCTCGTGTGGCTCGCTCGCTCTGTACTAGCCCTTAAGCAATTAAGACGAT 801
 Oy 693 gccagatatgatatcaaggaattgttggcaaaaaaatctcgtataagaccagatt 752
 Db 802 ACCAGATCATCTATTAGGACTTATTGAGAGCAAAAGATTTCTCCAGAGTGGCT 861
 Oy 753 tctcaagaactctgttatttaaccttggcagagtgatctctgacatgatttgaagtaa 812
 Db 862 TTGAAAGTGGCTGGATGCCACGCTTGCACTCATGTATCTGAAGAGAGCTGTGGAAA 921
 Oy 813 taccatgttactctcgtgtgattccaacacaaatataagaactgaagccgaagatgt 872
 Db 922 TCTCTGTTTCTCTCTGCTGATGATTAAGAGAAATTTAAATTCCTAGAGTGATGT 981
 Oy 873 atatgctggcccaactcttctgtgaacatctgtgcaaaatattctaacctggagcag 932
 Db 982 ATATACACACATCTCTGCTGCTGAACTGTGCAAAACATGTACACTGGACCCAGGC 1041
 Oy 933 agtgaattctgtgaactccggcattgactggtggagagtgagaccgaactgga 992
 Db 1042 TGTTAATTTCCAAAGCTTTTAAGCCTTTGACTGGGGAAGCAGTCCAGAAATTTTCA 1101
 Oy 993 atgcaatcagccaactctctgtaagtgatacagaagtcagaatgacggttccataaga 1052
 Db 1102 TTACAAACCAAGATTAATCTCCACATACATGTAAGAGACATGCTTGGCCGACGT 1161
 Oy 1053 gtgacagagaggtcagagactgcttcaaatccagaagacgtgaaatgtgtctctga 1112
 Db 1162 CTGAGAGGGGGTACACATGCTGCTGCAATGCTACAGAGTCAAAATCTTACTGACTCA 1221
 Oy 1113 ggtgaccaactcatctaccataagaatatctcgtgaatggcctgacgtgattcatctg 1172
 Db 1222 GATCACCAACTTGTGTTCCATGAGACATTCGGAATGGAGACATTTACTCTCATTTG 1281
 Oy 1173 gggcttgatgtctcctcaacgctatgtcaaatgaatcaatccatctgtgacagcag 1232
 Db 1282 GGGCTGTGATGCCCTTGAGAGCTTTATATAAATTAATTAATCTATGAGAAATATCA 1341
 Oy 1233 gaccacact 1241
 Db 1342 GTGAAGCT 1350
 RESULT 4

	HOMLIPCHL	2493 bp	mRNA	PRI	07-JAN-1995
	LOCUS	Human lysosomal acid lipase/cholesteryl esterase mRNA, complete cds.			
	DEFINITION				
	ACCESSION	M74775			
	VERSION	M74775.1			
	KEYWORDS	Lysosomal acid lipase/cholesterol ester.			
	SOURCE	Homo sapiens fibroblast CDNA from mRNA.			
	ORGANISM	Homo sapiens			
	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
	TITLE	1 (bases 1 to 2493)			
JOURNAL		Anderson, R.A. and Sando, G.N.			
MEDLINE		Clining and expression of cDNA encoding human lysosomal acid			
FEATURES		lipase/cholesterlyl ester hydrolase. Similarities to gastric and			
SOURCE		lingual lipases			
	J. Biol. Chem.	266 (33), 22479-22484 (1991)			
	920421292				
gene		Location/Qualifiers			
		1..2493			
		/organism="Homo sapiens"			
		/db_xref="taxon:9606"			
		/tissue_type="fibroblast"			
		/map="10"			
		<1..2493			
		/gene="LIPA"			
		/note="G00-120-153"			
		1..2493			
		/gene="LIPA"			
		41..1240			
		/gene="LIPA"			
		/EC_number="3.1.1.13"			
		/note="alternate initiator ATG may be at position 47"			
		/codon_start=1			
		/db_xref="GDB:G00-120-153"			
		/product="lysosomal acid lipase/cholesteryl esterase"			
		/protein_id="AAAS9519.1"			
		/db_xref="GI:187152"			
		/translation="MKMRFLGLVVCVLWPLHSBGGSKLTAVDPETNMNVSIIISYWGPEEELIVTEDGYILCLNIPIGRKNHSDKPKRPVFLOHGILLADSSMWTLMANSWSIGFTIADAGDVMGNRGTMTSRKRKTLSGDDEFAFSYDEMAYDLPASINFILIRNKGOEOVVYVGSGCGTTCIGTIAISOJPELAARIKMPCFALGPVASVAFCSPMAKLGRLPDHLIDLFQDGKEFLPOSATFLKMLGHVCTHYLIKLCELGLCFLICGFENRNLRSLVDVTTHSPAGTSVNMLHMGSQAVKFQAFQFDMSGSAKNFEHTNOSYPPTYNVKDMLVPRTVMSGGHDMLADVVDVINLLITQLTNLVFHESIIPMEHLDFTMGDAWRLYNKKIII NIMRKYO"			
		41..103			
		/gene="LIPA"			
		/note="alternate signal peptide may begin at position 47; as predicted from von Heijne, G. (1986) Nucl. Acids Res. 14,4683-4690; G00-120-153"			
		104..1237			
		/gene="LIPA"			
		/EC_number="3.1.1.13"			
		/note="G00-120-153"			
		/product="lysosomal acid lipase/cholesteryl esterase"			
		2466..2471			
		/gene="LIPA"			
		/note="G00-120-153"			
		2493			
		/gene="LIPA"			
		/note="G00-120-153"			
		BASE COUNT	692 A	484 C	516 G
		ORIGIN	801 T		
		Query Match	41.5%; Score 526.2; DB 9; Length 2493;		
		Best Local Similarity	65.3%; Pred. No. 2.2e-146;		
		Matches 790; Conservative	0; Mismatches 413; Indels	6; Gaps	1;
39	cagaatgaatatgtgcttcgatcattcgtagtgacgcgtatagtccagagaatgtgaattc	98			
07					
37	CAGAATGAATAAACCAGTTCTTGCGGGTGTGTGCATTGCTTGGTTGCCCTCCCATTC	96			

OY	99	ag-----tacatcgcacaataagctctgtggaccacagaagatccatgaatlatatgta	152
Db	97	TGAAGGGGCTTGAGAGGAAACTGTGACACTCTGTGATCTGTGAACAAACATGAATGTAGTGA	156
OY	153	aatatccaatatcaaaagcatccctcgtttaggaataatgacgcgaactgaaatggta	212
Db	157	AATTATCTTTACGGGGGATTCCTCGTGTAGGAATACCGTGTGAGACAGAAGATGGATA	216
OY	213	tatccttctgttaacaaaggttcctcgaggcctagtgcacaactaaagaacagtlccag	272
Db	217	TATCTGTGGCTTAACCGAATTCCTCATGTGGAGGAGAAACCATTCCTGCAAGGTCCCAA	276
OY	273	gccctgtgttactgtgaagatgacctagtgtgaggtgtgaacacttgatlttccaact	332
Db	277	ACCGATTGTCTCTCGAACATAGGCTTGGCTGGCAGAGATTCTTAATACGTGGTCACAAACT	336
OY	333	gcccaacaatagcctgtggtctcaatctcggcagaatgtgtttcttgacgttgatgggaa	392
Db	337	TGCCAACGACGAGCCTGGGGCTCATCTCTGTGTATGCTGGTTTGAACGTGTGGATGGCAA	396
OY	393	cagcaggggaaacgcctgtgtcctcgaaacaacaagacacatctccaataagaatgagt	452
Db	397	CAGACAGAGAAATACCTGGCTCTCGGAAACATAAGAACACTCTCAGTTCTCGAGATGATT	456
OY	453	ctggactttagttaataatagaatgctagtgttgaacctctcctgacgtgataaacttat	512
Db	457	CTGGGCTTTCAGTATATATAGATGGCAAAATATGACCTACCAAGCTTCCATTAACTCAT	516
OY	513	tttgcagaanaacgaggccaagaaaagatactatattgtcggctatctacagggcacaccat	572
Db	517	TCTGAATTAACCTGGCCAAAGACAAAGTATATTATGTGGGTCAATCTCAAGGACACATAT	576
OY	573	gggcttatgtcatlttcacacatgccaagcgtgtgctgcgaataatccaatgtatlttgc	632
Db	577	AGGTTTATACATTTCACAGATCCCTGAGCGCTAAAGATTAATAATGTTTTTGGC	636
OY	633	tttagcaccctatgcaactttaaagcatgcaaaaagcccgagacaacttttgtcct	692
Db	637	CCTGGGCTCTGGCTCCGTCCGCTTCCTGTACTACCTCATAGGCCAATTATAGACGATT	696
OY	693	gccagatatgatacaaggatgttttgcgaaaaaaagaattctgtlatcagaccagatt	752
Db	697	ACCAATCATCTCATTAAGGACTTATTGTGAGCAACAAGATTCTTCCCAAGAGTCCGT	756
OY	753	lctcagaacactgttatctaactcttgtgcccagcgtgacttctgtacagatttgaagaa	812
Db	757	TTTGAAGTGGCTGGGTACCCACGTTTGCACTCATATGCAATCTGAAAGAGCTCTGTGGAAA	816
OY	813	tatcatgttactctgtgtgatltcaacaacaaataatgaaacatgagccggaacagt	872
Db	817	TCTCGTCTTTCTCTGTGTGATTAATAGAAATAATTAATATGTCTAGAGTGTGATGT	876
OY	873	ataatgtccccaacactctgtcgtgaacatctgtgcaaaatatctaaccttggaccagc	932
Db	877	ATATCAACACATTCCTCTGCTGGAATCTTGTGCAAAACATGTTAACACTGAGGACAGC	936
OY	933	agtgaattctgtgaactccggcgaatttgcctgtgggagatgtgagcccaaaaacttgaaaa	992
Db	937	TGTTAATTTCCAAAAGTTTCAAGCTTTGACTGGGAGACAGTGCACAGAAATTAATTCCA	996
OY	993	atgcatacagccaactcctgtlaagatgcaagtcagagatalgacgyltccctacagcaat	1052
Db	997	TTACACACAGAGTTATCTCTCCACATCATATGTGAAGGACATAGCTTGTCCGACGTGACGT	1056
OY	1053	gtgcacaggaagtcagaagctgtgttcaaatccaagaagacgtgaaatgtcgtctctga	1112
Db	1057	CTGGAGCCGGGGGTACACATCGGCTGTGGAGATGTCTACGACGTCATATCTTACTACATCA	1116
OY	1113	gttgcacaactctatccacataagaaatctcctgaatgtggctcagctgattatcatctg	1172
Db	1117	GATCAACCACTTGTTGTTCCATGAGAGCATTCGCCGAATGGAGACATCTTGACTTCATTTG	1176

1173 99gttgatgctcctaccgcatatgacatgaatcaccatcctatcagcaggagga 1232
1177 GGGCCGTGATGCCCTTGAGGCTTTATATAAATTATTATATGAGGAATATCA 1236
1233 gaccacact 1241
1237 GTGAAGCT 1245

RESULT 5
BC012287 2586 bp mRNA PRI 08-AUG-2001
LOCUS BC012287
DEFINITION Homo sapiens, similar to lipase A, lysosomal acid, cholesterol
esterase (Wojman disease), clone MGC:5229 IMAGE:2900168, mRNA,
complete cds.
ACCESSION BC012287
VERSION BC012287.1 GI:15126726
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 2586)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (08-AUG-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
COMMENT Contact: MGC help desk
Email: cga@bbs.fda.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalob@bcm.tmc.edu
Villalob, D.K., Luna, R.A., Hale, S.M., Huiyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.

FEATURES
Source
1..2586
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:5229 IMAGE:2900168"
/tissue_type="Placenta, choriochorionoma"
/clone_lib="NIK-MGC_10"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
101..1300
/codon_start=1
/product="Similar to lipase A, lysosomal acid, cholesterol
esterase (Wojman disease)"
/protein_id="AAH12287.1"
/db_xref="GI:15126727"
/translation="MKRFLGLVCLVLPPLHSEGGKLTALDPENNMNVSEIISW
GFSESEYLVETEDGYLLCLNRLPHGRKNSDKPKPVFLDHLADSSMYNLIANS
SLGFLIADGPDVWNGNSRGTMSRKHTLSVQDFPAFSTDEMAKYDIPASINFL
NKTGEOVYVYVHSGSTTGFTLFSQIPELAKIKMFALGPVAVSACFSPAKLGR
LDPHLKIDLGDEKFLPOSAPFLKWLGTHTVILKELGKGLFLCGFERMLNMSR
VDVYTTSHSAGTSVONMLHMSOAVKFOKFAFPMGSSAKNYFNYSPTPVKML
VPTAVSGGHDLADYDVNLTITNLTNLFHESIPEMHLDPIWCLDAPMRLNKNII
NLMRKYQ"

BASE COUNT 736 a 511 c 531 g 808 t

ORIGIN
Query Match 41.3%; Score 524.6; DB 9; Length 2586;
Best Local Similarity 65.3%; Pred. No. 6.7e-146;
Matches 789; Conservative 0; Mismatches 414; Indels 6; Gaps 1;

39 cagaatgaaatgtgcttcgtatctggtggtggtatgcttcagagaatgltgattc 98
97 CAGAAATGAAATGCGGCTTTTGAGGTTGGTGGTCTTTGTTCTTGTGCCCCCTGCATTG 156
99 ag-----tacaatgccaactaaagtgtggaaccagaagaatcattgaatagta 152
157 TGAGGGGCTGTGAGGGAACCTGACAGCTTGGATCCTGAACAAACATGAATGATGATGA 216
153 aatcatcgaactcaaggtatccctgtgaggaatatgaatgcgaactgaaatgagta 212
217 AATTATCTCTTACTGGGATTCCTAGTGAAGAAATACCTAGTGGACAGAAAGATGATA 276
213 tatccttcctgtaacaggaatccctcgaagcctagtgcaactaaagaagaagttccag 272
277 TATCTGTCCTTAACCGAATTCCTCATGAGGAGAAACCAATTCGACAAAGTCCCA 336
273 gccctgtgtgtactgcaagcagcagcagcagcagcagcagcagcagcagcagcagc 332
337 ACCAGTGTCTCTCTGCAACATGCGCTTGTGCGACAGATTCCTAGTACCTGGGTCACAA 396
333 gcccaaatagcctcggggtcttaccctgagagatgctgagcagcagcagcagcagcag 392
397 TGCCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 456
393 cagcaggggaaacgcctgctcgaagaacacacacacacacacacacacacacacacac 452
457 CAGCAGAGGAATTAACCTGCTCGGAAACATTAAGACATCTGATCTTCACAGATGAAT 516
453 ctggccttcagatgataatgataatgataatgataatgataatgataatgataatgata 512
517 CTGGCTTTCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 576
513 ttgagaagaacagcggcgaagaagaatctatgattgcttaccagggcagcagcagcagc 572
577 TCTGAATAAACCTGCGCAAGAACAGATATATATGATGATGATGATGATGATGATGAT 636
573 ggccttattgcatcttcac 632
637 AGGTTTATACATTTTACAGATTCCTGAGTGGCTTAAAGATTTAAATGATTTTGGC 696
633 ttgac 692
697 CCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 756
693 gccagatatgataatgataatgataatgataatgataatgataatgataatgataatg 752
757 ACCAGATCATCTCTCTGAGGATTTTGAAGAACAAATTTTCCCGCAGAGTCTT 816
753 tctcagaac 812
817 TTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 876
813 tacaatgcttactctgtgtggtatcacaacacacacacacacacacacacacacacac 872
877 TCTCTGTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 936
873 ataagcgcgcac 932
937 ATATACAAACATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 996
933 agtgaattcgtgtgaactccgagcattgagcagcagcagcagcagcagcagcagcagc 992
997 TGTTAATTCACAAAGTTTCAAGCCTTGGACTGGGGAAGCAGTCCCAAGAAATTTTTC 1056
993 atgcaatcagcacaactcctgtgaaggtacagagtcagagatagcaggtccctacagca 1052

Db	496	ATGACCTTCACAGCACAATAATTGACTTATTTGAAGAAACTGCAGAGAAAGCTGCAGCT	555
Qy	545	atgcgcgtattcaagagcgagccacacatcaggcttattgcatttccaccatgccagagc	604
Db	556	ATGTGGTGTCAATCCACAGGGGACCCACCATTTGGTTTATTTCCTCTTCTCAATCCCAAGC	615
Qy	605	tggctcgaataatcaaatgatttcttgcttgagaccctaaagcactgattgaagctgcaa	664
Db	616	TGGCTGAAAGAAATCAAAACCTTTTATGCAATTAGCTTCAGTCCGCACTGTGAAGTATACAA	675
Qy	665	aaagcccccggagccaaattttctgcgcagatatgataatgatacaaggaattgttgcga	724
Db	676	AAACCCCTGTGAACAAACAACTAGGTATTATTCCTCCCAACATGTTCAAGATTAATTTGGTG	735
Qy	725	aaaaagaattctgatalcagacaggaattctcagacaactgttatttaaccttgtgc	784
Db	736	ACAAAAATATTTCTACCCACAAATTTCTTGATCAATTTCTCCACCTCAAGTGCTGCC	795
Qy	785	agggatattcttgatcagattgtagaataatcaltgltacttcttggttgattcaacca	844
Db	796	GTGGAGACACTGAAGTAAATTGGACGAATGCCCTTATTTATTCATTTGTGGCTTGACACGG	855
Qy	845	acaatatgaacaatgagccgcgagcaagtatcaltgcgccacactctgtgaaacatctg	904
Db	856	CAAACTTGACATGATGATCGCTTGATGATGTACGTGCACATTAATCCGGCAGAACTTACG	915
Qy	905	tgcacaaatattcacaactcagcagcagcagtgaaattctgtgtaactcggcgattgact	964
Db	916	TTCAAAACATGCTGCACATGTGACCCAGCGTGTAAATCTGGAATTTTCAAGCTTTAATT	975
Qy	965	ggggagattgagacccaataatctgnaaaatgcaatcagcacaactccgtlaagltacagag	1024
Db	976	GGGGAAGTCAGACTCAAAATGATGTGCTACATCATCAGCCCACTCCCTCTACAAATG	1035
Qy	1025	tcagagatatgagcgtctcccaagcaatgtgagcaggaagtcgaagcgtcttcaaatc	1084
Db	1036	TGACCCGCATGAATGTCCAATTCGAGTGGGAGTGTGGCAATGATCGTGGTGTGAC	1095
Qy	1085	cagagacagctgaaaatcgtctctcctcgtgagtgagccaaactcattacataagaatctc	1144
Db	1096	CCCAAGATGTTGACCTTTTGCTTCCCAAACTTCTTAATCTCATTTTACCAAGAGGATTC	1155
Qy	1145	ctgaatgggctcagctgagatttcaatcctggggtttggaatgctcccaactgatacgaatg	1204
Db	1156	TTCCATTAATCACTTGATTTATTCGTGGCAATGATCTCCTCAAGAAATTTTCAATG	1215
Qy	1205	aaatcatcactgtgatgacagcagggagaga	1234
Db	1216	AAATTAATTTCTATGATGCAAAAGATTAATA	1245
RESULT	7		
LOCUS	A26690	1378 bp	DNA
DEFINITION	Precursor of rabbit gastric lipase (comp.).		PAT
ACCESSION	A26690		
VERSION	A26690.1	GI:905030	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS	1 (bases 1 to 1378)		
TITLE	Benicourt,C., Blanchard,C. and Junien,J.L.		
JOURNAL	Recombinant gastric lipase from rabbit and pharmaceutical		
FEATURES			
source			
BASE COUNT	418 a	276 c	287 g
ORIGIN			

Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;			
QY	119	ctgagaccgaagcattcatgaattatgtaaatcatccacatgaagctatccct	178
DB	121	CTGGAACCCCTGAAGTACTGATGAACATTAATGATGATTAATGAGGATCCCAA	180
QY	179	gtgaggaatgaagtcgaactgaagaatggatattccttcgttacaagattcttc	238
DB	181	ATGAAGATATGAAGTGTGATGCTGAAGTGTATATCTTGAAGTCAATGAATTCCTT	240
QY	239	gaagcctagtagcaactgaagaagcaggttcacagcctggtgttactgaagcatggcc	298
DB	241	ATGGAGAAGAAAATTGACGGGAATACAGGCCAGAGACCTGTGTGTTTTCGACGATGTT	300
QY	299	taagttagagtgctagaacatggattcccaactgcccacaataagacttggctcatc	358
DB	301	TGCTTGATCATGACCAAACTGGATTTCACACCTGCCAACAGACGCTTGCTTCATTC	360
QY	359	tggagagtgctggttttgagtgatgggaaacagaggggaaacgctggtctgaa	418
DB	361	TGGAGATGCTGTTATGATGTGTGGTGGCAACAGAGGAAGAACCTGGGCCAGAA	420
QY	419	aacacaagacactcctcaatagaacaaatgagttcgtggtcttcagtatgagatgg	478
DB	421	GAACTGTATATATTCACACGATTCAGTGAATTCGTGGCTTTCAGTTGATGAATGG	480
QY	479	ctaggtttgaccttcctgagtgataaacttattttgcagaaaaacgggccaagaaaga	538
DB	481	CTAATATATGACCTTCACAGCCACATTCGACTTCATGTGAAGAAAACCTGGACGAGC	540
QY	539	tctattatgctgctatcaagaagggacacacacacacacacacacacacacacacac	598
DB	541	TACACTATGTTGGCCATCCACGAGGACACACCATGTTGTTATTCCTTCACCAATC	600
QY	599	cagagcggctcagaaaaatcaaatgtatttctgcttagcaccacacacacacacacac	658
DB	601	CCACCTGGCTAAAGATCAAAACCTCTATGCTCTAGCTCTGCTTGGCACTGGAGCT	660
QY	659	atgcaaaaagcccggaacaaatttctgtctgctgcagataatgatacaaggaatgt	718
DB	661	ATACAAAAGCCTTATAAACAAACTTGAATTTGTTCTCAATCCCTTCAGATTATAT	720
QY	719	ttggcaaaaagaattctgtatcagaccagaatttctcagacaaattgttattacctt	778
DB	721	TTGGTGAACAAATATTTCTACCCACACACTTCTTGATCAATTTCTCTACTGAGTGT	780
QY	779	gtggcagagttattcttgatcagatttgtaataatcatgttacttcgggtgatca	838
DB	781	GCTCCGTGAGATGCTGAATCTCTTTGCAGCAATGCTTATTTATTTTGTGATTTGG	840
QY	839	acaccacaatatgaacatgagccgagcaagtgtatattgtctgcccacactctgtgaa	898
DB	841	ACAGTAAAGAACTTAAACAGAGTGCCTTGATGTATCTATCAATATATCAGACAGAA	900
QY	899	catctgtgcaaatattctaacatgagccagcgagtgatctgtgttaaccctgggat	958
DB	901	CTTGTGTTCAAAACATGTTTCCATTTGAGCCAGCGCTTAAAGCTGGGAAATTCACAG	960
QY	959	ttgactggggagtgagacaaaatctggaataaattgcaatcagcaactcctgtgaagt	1018
DB	961	ATGACTGGGGAGCCCACTTGAATAGGATGACATATGATGATGCCAACCTCCCTACT	1020
QY	1019	acagagacagatatgacggttccctacagcaatgtgagcaaggaagtgacagctgtt	1078
DB	1021	ACATGTGACAGCCATGATGATACCAATTTGAGTGGAGCGTGGAAGGACGCTGTGG	1080
QY	1079	caaatccaagaagcgtgaataatgtctctctgaggtgacccaactcatcacaataga	1138
DB	1081	CTGACCCCAAGATGTTGGCTTTGCTTCACAAAACCTCCCAATTTTATTTTCCACAA	1140
QY	1139	atattcctgaatggcctgagttcatcctgaggttgcagttcctcctcaccglatgt	1198
DB	1141	AGATTCCCTTTTACATATCTTGACCTTATCTGGGCAATGGATGCCCTCTCAGAAAGTT	1200

QY	1199	acaatgaatcatcatcgtgat	1220
DB	1201	ACAAATGACATTTGTTCTATGAT	1222
RESULT 9			
LOCUS	A01046	1367 bp	mRNA
DEFINITION	H.sapiens mRNA for human gastric lipase.		
ACCESSION	A01046		
VERSION	A01046.1	GI:344241	
KEYWORDS	gastric lipase.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 1367)		
TITLE	POLYPEPTIDE AND POLYPEPTIDE COMPOSITION		
JOURNAL	Patent; WO 8601532-A 5 13-MAR-1986;		
FEATURES	Location/Qualifiers		
source	1..1367		
sig_peptide	/organism="Homo sapiens"		
CDS	/db_xref="taxon:9606"		
	47..103		
	/note="gastric lipase"		
	/codon_start=1		
	/product="lipase"		
	/protein_id="CAA00125.1"		
	/db_xref="GI:344242"		
	/db_xref="SWISS-PROT:P07098"		
	/translation="MWLLTMASLSVLTGTHGLEFKLIHPSPVEYFNIISQNTIYWGYPNEEVEVTEDEGYLIEVNRIPYGRKNSGTQRPVYFQHGILASATWMSISLPPNNSLAFILADYDYLWLGSRGTWARRNLYSPSPVEWAFSPDEMAKYDIPATIDFVYK		
	TGOKDLHYVGHSGDTGIEFIAFSTNPISLAKRIKPYALAPVATYKSLINKLEFVY		
	OSLRFITGDEITFPHNFDFPLAREVSRREMLNCSNALEFICGSPSKNPNSTRD		
	VYLSNPFATISVONFHTQAVKSKRFATDMGSGFVQRMHRTDOSQPPYINTANVP		
	IAVWNGKDLADPDVGLLPLKPLNLYHEIPYNNHLDFTMADAPQEVYINDIVSM		
	ISEDK"		
	104..1240		
mat_peptide	/note="gastric lipase"		
	/product="lipase"		
BASE COUNT	397 a	284 c	269 g
ORIGIN	417 t		
Query Match			
Best Local Similarity 64.2%; Pred. No. 1.4e-129;			
Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;			
QY	119	ctgtgagaccgaagcattcatgaattatgtaaatcatccacatgaagctatccct	178
DB	123	CTGGAAGCCCTGAAGTACTGATGAACATTAATGATGATTAATGAGGATCCCAA	182
QY	179	gtgaggaatgaatgaactggaagaatggatattccttcgttacaagattcttc	238
DB	183	ATGAAGATATGAAGTGTGATGCTGAAGTGTATATCTTGAAGTCAATGAATTCCTT	242
QY	239	gaagcctagtagcaactgaagaagcaggttcacagcctggtgttactgaagcatggcc	298
DB	243	ATGGGAAGAAAATTGACGGGAATACAGGCCAGAGACCTGTGTGTTTGCAGCAATGTT	302
QY	299	taagttagagtgctagcaactgatttcaactgcccacaataagcctgggtctcatc	358
DB	303	TGCTTGATCATGACCAAACTGGATTTCACACTGCGCAACAAACCTTGCTTATTC	362
QY	359	tggcagatgctggttttgacgctgtgagatgaggaacagcaggggaaacgctggtctgaa	418
DB	363	TGGCAGATCTGCTTATGATGTGTGGCTGGCAACAGCAGAGGAACACCTGGGCCAGAA	422
QY	419	aacacaagacactcctcaatagaacaaatgagttcgtggtcttcagttatgagatgg	478

Db 423 GAACTGTACTATTCACAGATTCAGTGAATCTTGCGCTTCAGCTTATGAAATGG 482
 Qy 479 ctgaagttgacctccgcgcagtgataaacttatttgcagaanaacggcgcaaganaaga 538
 Db 483 CTAATATATGACCTTCAGCCCAATGCAATCTTATGTAAGAAACCTGACAGAAAGCAGC 542
 Qy 539 tctatttgcgctattcacaagggaaccaccatggccttattgcatcttccaccatgc 598
 Db 543 TACACTATGTGGCCATTCCAGGCGACACCATTTGGTTTATTTGCTTTTCACACATC 602
 Qy 599 cagagctgctcagaanaatacgaatgtatttgcctttagcaccatagccacgttgaac 658
 Db 603 CCAGCTGGCTAAAGATATCAAAACCTCTATGCTCTAGCTCCGTTGCCACTGTGAAGT 662
 Qy 659 atgcaaaaagcccgaggacaacatttctgtctgcagagataatgatacgaaggattgt 718
 Db 663 ATACAAAAGCCCTTATTAACCAACTAGATTGTCTCAATCCCTCTTCAAGTTTATAT 722
 Qy 719 ttggcaaaaagaattctgtatcagaccagattctcagaacaactgttattaccctt 778
 Db 723 TTGGTGCACAAATATTTCTACCCACACAACTTTTGATCAATTTCTGCTACTGAAGTGT 782
 Qy 779 gtggcaggtgattcttgcatacagaattgttagtaataatcattgcttgcgtgattca 838
 Db 783 GCTCCCGTAGATCTGATCTCTTCTTGACAGCATGCTTATATATTTGTGTGATTG 842
 Qy 839 acacacaataatgatacagatgagccgagcaagtgtatgtctgcacaaactctgtcgaa 898
 Db 843 ACACTAAGAACTTACACAGAGTCTGCTGATGTCTATCTATCAATATACACAGAA 902
 Qy 899 catctgcaaaaataattctcactgagcgagcgaggaatcttcgtggaactccggagat 958
 Db 903 CTTTGTGTCACAAATATCTTCATGTGACCCAGGCTGTAACTCTGGGAAATTCACAGCTT 962
 Qy 959 ttgacttgaggagagagaccaaactctgaaactgcaacacacacacacacacacacac 1018
 Db 963 ATGACTGGGAGAGCCGACATTCAGATAGATGACATGACATGACATGACATGACATGAC 1022
 Qy 1019 acagagtcagagataatgacggtccctcagaacatgtgacagagagagacacacacac 1078
 Db 1023 ACATATGACAGAGCATTAATGACATGACATGACATGACATGACATGACATGACATGAC 1082
 Qy 1079 caaatccagaagaacgtgaaatgtctgtctcctgagtgagacacacacacacacacac 1138
 Db 1083 CTGACCCCAAGATGTGGCTTTGCTTCCAAACCTCCCAATCTATTTACCAACAGG 1142
 Qy 1139 atattcctgaatgggtcagctcagatctcattcctgaggttgagatgctcctcaccgtatgt 1198
 Db 1143 AGATTCTTTTACATCACTTGACCTTTATCTGGCAATGATGCCCCCTTCAGAGATT 1202
 Qy 1199 acaatgaatcaccatcattgat 1220
 Db 1203 ACAATGACATTGTTCTATGAT 1224
 RESULT 10
 A12714 1367 bp DNA PAT 09-AUG-1994
 LOCUS A12714
 DEFINITION Pregnastic lipase.
 ACCESSION A12714
 VERSION A12714.1 GI:579560
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1367)
 AUTHORS PRECURSOR POLYPEPTIDE, DNA SEQUENCE CODING THEREFOR, VECTORS, HOST
 TITLE ORGANISMS, AND PROCESSES INVOLVING SAME
 JOURNAL Patent: WO 8603778-A 8 03-JUL-1986.
 FEATURES Location/Qualifiers

source 1..1367
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 CDS 47..1243
 /codon_start=1
 /product="pregnastic lipase"
 /protein_id="CAA01053.1"
 /db_xref="GI:579561"
 /db_xref="SWISS-PROT:P07098"
 /translation="MMLLTMASLISVLGTHGLFGKLPHGSEPTWNISQMITWGY
 PNEYEVYTEDGYILEVNRIPYKKNKNGNGORVPELQGLASATNMTSNPNNSL
 AFLIADGYVWMLGNSRGNWARMNLYSDSVEAFSEDMAXYDLPATIFEVYK
 TGOXOLHYVHSGSTTGIFLAFSPNPLARITFYALAVATVYKYSILNLRFP
 QSLPKFTGDKITYPHNFEPQPLATEVCESEMNILCSNMLFICGDSKNPTSRD
 VYLSHNPAGTSVQNMWSPHVTQAVGSKQVYKIDWESPVONRHHYDOSPTYNVAMNP
 IAVNGSKDLADPODVGLLPKLPNLVYHKEIPYNNHDFIWMADAPDEYNDIVSM
 ISEDKK"
 BASE COUNT 397 a 284 c 269 g 417 t
 ORIGIN
 Query Match 37.0%; Score 470; DB 6; Length 1367;
 Best Local Similarity 64.2%; Pred. No. 1.4e-129;
 Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;
 Db 119 ctgtgagaccagaagcattcctgataatagtaataatcacaacatcagaagctatccct 178
 Db 123 CTGGAACCCCTGGAAGTACATATGAACTTACTGATGATTAATTTGGGATACCCAA 182
 Qy 179 gtggagataatgagtgacacatgaagatggtatcccttctgttaacagagatcttc 238
 Db 183 ATGAGATATGAACTTGTGCTGACATGAGTGTATATCTTGAAGTCAATGAAATCTCTT 242
 Qy 239 gaagcctlagtgcaacttaagaagaacaggttccagagcctgtgtgttactgcagatggcc 298
 Db 243 ATGGGAAGAAAATTCAGGGAATTCAGGACCCAGACACCGTGTGTTTTCAGCATGTT 302
 Qy 299 taattgagatgtcagacactgtgatctcacaactgcacacaatagcctggtctatc 358
 Db 303 TGCTTGATCAGCCCAACAACTGAGTATTCACCTGCGGCAACACAGCCTTCTCATATTC 362
 Qy 359 ttgagagatgtcgttttgagtgatgagtgaggaacagagggagaaagccgtgctcgaa 418
 Db 363 TGGCAGATGCTGTTATGATGTGTGCTGGCAACACAGAGAAACCTGGGCGCAGAA 422
 Qy 419 aacacaagaacactcctcacaagacagatgagtgcttggtcctcagttatgatatgagtg 478
 Db 423 GAACTGTACTATTCACAGATTCAGTGAATCTTGCGCTTCAGCTTATGAAATGG 482
 Qy 479 ctgaagttgacctccgcgcagtgataaacttatttgcagaanaacggcgcaaganaaga 538
 Db 483 CTAATATATGACCTTCAGCCCAATGCAATCTTATGTAAGAAACCTGACAGAAAGCAGC 542
 Qy 539 tctatttgcgctattcacaagggaaccaccatggccttattgcatcttccaccatgc 598
 Db 543 TACACTATGTGGCCATTCCAGGCGACACCATTTGGTTTATTTGCTTTTCACACATC 602
 Qy 599 cagagctgctcagaanaatacgaatgtatttgcctttagcaccatagccacgttgaac 658
 Db 603 CCAGCTGGCTAAAGATATCAAAACCTCTATGCTCTAGCTCCGTTGCCACTGTGAAGT 662
 Qy 659 atgcaaaaagcccgaggacaacatttctgtctgcagagataatgatacgaaggattgt 718
 Db 663 ATACAAAAGCCCTTATTAACCAACTAGATTGTCTCAATCCCTCTTCAAGTTTATAT 722
 Qy 719 ttggcaaaaagaattctgtatcagaccagattctcagaacaactgttattaccctt 778
 Db 723 TTGGTGCACAAATATTTCTACCCACACAACTTTTGATCAATTTCTGCTACTGAAGTGT 782
 Qy 779 gtggcaggtgattcttgcatacagaattgttagtaataatcattgcttgcgtgattca 838
 Db 783 GCTCCCGTAGATCTGATCTCTTCTTGACAGCATGCTTATATATTTGTGTGATTG 842

[illegible]

OY	599	cagagctggtcgcgaagaatacaaaatgatatttgctttagcacccaatagccactgtttaac	658
Db	560	CTAAGTGGGCAAAAAATATAAGCTTTTAAGCATTAAGTCACAGTGCGTACCCTGAATG	619
OY	659	atgaaaanaagccccggagccaatatlttgctgcccagatalbatgacaaggatgct	718
Db	620	ATACACAAGTCCCTTGAAAAGAATTTCATTATTTATCCTCACATTTGCTTTCAAGCTTAGT	679
OY	719	tttgcaaaaaagaatlctctgatcagaccagattlcttcagacaacttgttatccctt	778
Db	680	TTCGCACGAAGAAATGTTCCTCCCCACACACTTACTTATGATACCTTTCTTGGATACGAAGTGT	739
OY	779	gtggccaggtgatctctgcatcagatlttgttaataatcatagttacctcgggtgattica	838
Db	740	GCTACGGGAGGTTCTTGATCTTCTCTGCAGAACACTTATTACCTCTGTGGATTTTG	799
OY	839	acaccaacaatalgacalpagacpgccgagcaagtgtatatactgcgccacactcttgtgaa	898
Db	800	ACAAGAAAACCTTAAGTGTAGTCTGTTTTGATGTGTATGTAGGGCATATGCCAGCAGGGA	859
OY	899	catctgtgcaaatatctctcaactctgagccgaagcaatgtaattctgttgaaatccgggat	958
Db	860	CATCTGTTTCAGAGCTTCTCTCACTGGGCAACGCTTTTGATGTGGGAAATTTCAAAGCT	919
OY	959	ttagctgtggggaatlgagaccaaaatctggnaaaatgcaatcagccaactcctgtlaagt	1018
Db	920	TCAAGTGGGAAAGCCCATCCCAAGACATGTTAACTACATACACAGAAAGAGCCTCCTGAAT	979
OY	1019	acagaagtcagagalatyaagctccctacagcaaigtgascaggaaugtlcagsgactgcttt	1078
Db	980	ATGATGTGTCAGGCATACGTACTGTCGCCAGTTGCCAATGTGGAAACGGTCCAATTCATCCTGG	1033
OY	1079	caaatccagaagacgtgaaatatctgctctctgaggtlgacccaactcactcattacalaaga	1138
Db	1040	CTGATGCCCAAGATGTGCATATCTGCTTCCCAAACTCTCAACCTCCTGTTCCATTAAG	1099

Db 1100 AGATTCTTGCCCTACAATCACCCTGACTCACTCTGGGCAATGGATGCCCCCTCAGAGGTTT 1155

Db 1160 ACAATGAGATGATTTCCATGATGGCAGAGACTAAAGATTC 1201

RESULT 12

RNLIIP 1355 bp mRNA ROD 30-MAR-1995

LOCUS Rat mRNA for lingual lipase.

DEFINITION X02309.1 GI:56595

ACCESSION X02309.1 GI:56595

VERSION glycoprotein; lipase.

KEYWORDS Norway rat.

SOURCE Rattus norvegicus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 1355)

AUTHORS Docherty,A.J., Bodmer,M.W., Angal,S., Vargier,R., Riviere,C., Lowe,P.A., Lyons,A., Emtage,J.S. and Harris,T.J.

TITLE Molecular cloning and nucleotide sequence of rat lingual lipase cDNA

JOURNAL Nucleic Acids Res. 13 (6), 1891-1903 (1985)

MEDLINE 85215587

COMMENT Data kindly reviewed (12-FEB-1986) by A.J.P. Docherty.

FEATURES

Source Location/Qualifiers

1..1355

/organism="Rattus norvegicus"

/db_xref="taxon:10116"

49..102

/product="put. signal peptide (aa -18 to -1)"

49..1236

/note="lipase precursor"

/codon_start=1

/protein_id="CAA26179.1"

/db_xref="GI:56596"

/db_xref="SWISS-PROT:P04634"

/translation="MWLLIISVISTREGANGLGKLPENPMNISMHTTGP
CQETEVTEDEYILGVIRIPKNNSEIGRPVYVYLQHLIASATNMIANLPNNLSLA
FMLADAGYDVLGNSRGMTWRSKNVYSPDSVEWASFDPMARYDLPATINFTVOKT
GOEKIHVHGSGQGTIGFIASFTNPTLAKKKTVALAPVATVYTOPLKRIEIP
FLFKMGKSMFLPHYTFEDDELGTGECREVLGLCSNTLFTFCGPKKNLVSRP
YLGNPAGTSVODFLHNAOLYRSGKFOAFNMGSSONLHYNOKTPPYDVSATTPV
AVWNGNDIADPDVAMLLPKLSNLEFHKELILVNLHDFIMANDADQEVNENISM
AED"

mat_peptide 103..1233

/product="lipase (aa 1-377)"

misc_feature 505..507

/note="aa 153 pol. lipid binding site"

misc_feature 1355

/note="polya site"

BASE COUNT 400 a 291 c 276 g 387 t 1 others

ORIGIN

Query Match 36.8%; Score 467.6; DB 10; Length 1355;

Best Local Similarity 63.5%; Pred. No. 7.6e-128;

Matches 713; Conservative 0; Mismatches 409; Indels 0; Gaps 0;

Db 362 TAGCAGCGCTGGTATGATGTGGTGGGAACGTCGAGCAATATCATGTCCACGA 421

419 aaccagaagacactccatagacaagaatgagttcgtggcttcagttatgagatg 478

422 AAAATGTACTACTACACAGACCTCATGTAATTTGTGGCTTTACGCTTTGTAATGG 481

479 ctaggcttgacccctccctcagtgataaaccttatttcagagaagaagccagagaaga 538

482 CTAAATATGACTTCCCGCCACAAATRAACTTATTGTACAGAAATCGACAGAGAAGA 541

539 tctattatgctggctattcaacaggcaccacacatgaggtcttattgcatllccaccatgc 598

542 TACACTATGTGGTGTCTCTCCAGGGCACCATATGTGTTTCCTCTTCTTACCAATC 601

599 cagactggtcctcagaataatcaaatgatttttgccttgagaccatcagcactgttaagc 658

602 CTACACTGGCCAAAAAATCAAGACGTTTATGCAATTAAGCTCCAGTTCCTACCGGAAGT 661

659 atgcaaaaagcccgagccaattttgtgctgcagatatgatacgaaggatgt 718

662 ATACCAAAAGTCCCTTGAAAAAGATTATTTCTACATTTCTTTCAAGCTTAATGT 721

719 ttggcaaaaagaattctgatacagaccagaattctcagaacaattgattaccttc 778

722 TTGGCAAGAAATGTTCTCCGCCACACCTACTTGTATGATGATTTCTTGTAACCAAGTGT 781

779 gtggccaggtattcttgatcagatttgagtaataatgatttacttctgggtgattca 838

782 GCTACGGGAGGTTCTTAATCTTCTCTCAGACMACCTTATTCATCTTCTGTGATTTGG 841

839 acaccaacaatatgacatgagccgagcaagtgatagatgctggccaccactctgttgaa 898

842 ACAAAATAAATTAATATGAGTCTTTTGATGTATTAATCTAGGCGATTAATCAGCAGGGA 901

899 catctgtgcaaaatattctacacttgagagccagcagtgatcttggtgaactccggagcat 958

902 CATCTGTTCAAGACTTTCTCCAGCTGGGACGCTGTTAGATTCGGAAATTTCAAGCTT 961

959 ttgactggggagtgagagccaaaataatctgggaataatgcaatcagccaactctctgaaagt 1018

962 TCAACTGGGGAAGCCCATCCAGAACATGTATACATTAACCAAGAAAGCGCTCTCGAAT 1021

1019 aagaagtcagagatatgacggtgccctacagcaaatgtgacagaagtgacagctgtt 1078

1022 ATGATGTCTACGCCATGACTGTGCCAGTGTGCAAGTGGCAATGCAATGACATCCTGG 1081

1079 caaatccagaagacgtgaaaatgctctctctgagtgagtgacaaacctcatcataaaga 1138

1082 CTGATCCCCCAAGATGTCCGCAATGCTGCTCCCAAACTCTCCAACTCCTGTTCCATAAGG 1141

1139 atattccgaatgggtcagctgagttcatctgagtggtttgagatgctctcaccgtatgt 1198

1142 AGATTCTTGCTACATATCCTGGACTTCATCTGAGGCAATGAGATGCCCTCAAGAGTGT 1201

1199 acaatgaaatcatcatctgatactgagcagcagagggaggaaccaacc 1240

1202 ACAATGAGATGATTTCCATGATGAGGAGCAAGCAATTAANAATTC 1243

RESULT 13

MLLYACLY 2358 bp mRNA ROD 24-FEB-1995

LOCUS Mmusculus (C57 Black/6J CBA) LAL mRNA for lysosomal acid lipase.

DEFINITION 231689

ACCESSION 231689.1 GI:4456670

VERSION LAL; lipase; lysosomal acid lipase.

KEYWORDS house mouse.

SOURCE Mus musculus.

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2358)

AUTHORS Du,H., Witte,D.P. and Grabowski,G.A.

JOURNAL Mutation in the rat model of Wolman's disease
MEDLINE J. Lipid Res. 36 (10), 2212-2218 (1995)
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gidsq 176569] from the original journal article.
This sequence comes from Fig. 2.

FEATURES

Source

1. 3144
/organism="Rattus sp."
/db_xref="taxon:10118"

gene

1. 3144
/gene="lysosomal acid lipase, LAL"

CDS

233..1426
/gene="lysosomal acid lipase, LAL"

/note="intracellular hydrolase; This sequence comes from
Fig. 2; conceptual translation presented here differs from
translation in publication; LAL"

/product="lysosomal acid lipase"
/protein_id="AAB36043.2"

/db_xref="GI:9653291"

/translation="MQLGRVLCFVVGILLSGPTGTISAVPEANNMVEIIMHMGY
PEHSVONGDGLILGHRIPGRKNOPKPRPYVQLMRHGFELADSSMWTNINSL
GFTLADAGFDWMGNSRGNMRSRKHTLSVODPYMPSPEMAKXDLPAIYTLNK
TGQKOLNVGHSGCCTGTFTASOMPELAKVKRFFALAPVLSLNFSGPVKLRIP
DLLEDLFGKQKPLPQSAVMKWLSTHCTHVMELCANIFELIGENKLNMSRD
VYTHCPAGTSVQNMVHTQVYKHKLOAPMGSDKNRPHYNGSYPLYSIKMOLP
TALMSGKDWLADSDINILLETIPLYVHKNIPEMDHIDEIMGLDAPWRLYNEVSL
MKRYQ"

BASE COUNT 852 a 690 c 754 g 848 t
ORIGIN

Query Match 36.2%; Score 459.8; DB 10; Length 3144;
Best Local Similarity 64.5%; Pred. No. 1.9e-126;

Matches 724; Conservative 0; Mismatches 387; Indels 12; Gaps 2;

OY 117 agctgtgaccagaagcattcattgataattgtaaatcatcaacatcaagctatcc 176
DB 307 agctgtgaccagaagcattcattgataattgtaaatcatcaacatcaagctatcc 366
OY 177 ctgtgaggaataatgagctcgcaactgaatggtatattcccttctgttaacagattcc 236
DB 367 a-----gagcactcagatracagacagagatgcttacccttctgttaccgaattcc 420
OY 237 tctgagcctagtgtcaacttaagaacaggttccagcgtgtgtgttactgtcag----- 291
DB 421 gcattggcggaagacacagttgattgaaaggtcccaaacagttgttcttcttctgagtgccg 480
OY 292 -catgctagtgtgaggtgtcagcaactgatttccaaactgtcccaacaataagcttgg 350
DB 481 acatggcttcttgcgacagatttctgactggtcacaacacattgacacacagctggcg 540
OY 351 cttaattctgagagatgctgtgttctgactgtgtgagtggtggaacacaggggaaagcctg 410
DB 541 ctctcatctctgacagatgctgcttctgattgtgtgagtggaacacacagggaaacacctg 600
OY 411 gttctgaaacacacagacacttccatagacacagatgagttctgtggtcttcaattatga 470
DB 601 gttctgaaacacacagacacttctcagatttctcagatgagttctgtggtcttcaattatga 660
OY 471 tgaagtgtgagttgttgaacttctcgtcagtgatataacttatttgcagaaacagggcca 530
DB 661 tgaagtgtgagttgttgaacttctcgtcagtgatataacttatttgcagaaacagggcca 720
OY 531 ggaagaatctatgtctgagctattcagagggcacacacatgggcttattgacatttc 590
DB 721 agaacagctgtatgtgtgagcactctcagagcctgacacatagctttttatgacatttcc 780
OY 591 caaccatgcagagctgtgtcagaaacacaaatgtatttctgttgcacacatagccac 650
DB 781 acagatgctcgttaagcgaagacagattgatttctgttgcacacatagccac 840

OY 651 tgttaagcatcaaaaagccccgggaccacaattttgtctgtccagatgatgatcaa 710
DB 841 tctcaattttcttccgcccgcagatgctcaaatgtggagcaggtgctggcgttcttga 900
OY 711 gggattgttgcacaaagaattctgtatcagacagcagattctcacaacattgtat 770
DB 901 ggacttgtttggcgacgaagacatttcttccctcagagtcatttggaatggtctagcac 960
OY 771 ttaccttgtgtccaggtgattctgtatcagaattgttgaataatcagatgttacttctgg 830
DB 961 tcacatttgcacagcagctcattcattcagagagcttggcccaacattcttcttcatatg 1020
OY 831 tggattcaacacacaaataatgaaatgagcagcaggtgtatattgtgtcccaactct 890
DB 1021 tggcttcatagaaagattttaaatacattgtgagctgagtgatgtgtatcacaacactgtcc 1080
OY 891 tctgtgaacatctgtgcaaaaatattctcactgtgagcagcagcaggttctgtgtaact 950
DB 1081 tgcgggaacctgtgtgcaaaaacattgtcactgtgacccagctgtttaaatacattacact 1140
OY 951 ccggcatttgcactgtgagggagtgagaccaaataatctggaataatcagcacaactcc 1010
DB 1141 tcaacttcttgcactgtgagggagtgagaccaaataatcttcaattacattcagagttatcc 1200
OY 1011 tgaaggtacagagtcagagatataagcggctccctcagcaatgtgtgacagaggtcaga 1070
DB 1201 acccttgtacagcatttgaagaaacattcactcactgttatttgagacggggcgacaa 1260
OY 1071 ctggtcttcaatccagagaacgtgaaatgtgtctctcgtgagtgagcacaactcatcta 1130
DB 1261 ctggcttgacagacacacagcattcatttcttcttcttcttcttcttcttcttcttctt 1320
OY 1131 ccataagaatatctcgtgaatggtcagctcagctgatttcttgggttggatgtcctca 1190
DB 1321 ccacagaagaacattcccgagtgagacacacattgatttcttcttcttcttcttcttctt 1380
OY 1191 ccgtatgtacatgaaatcattcattcattcattcattcattcattcattcattcattc 1233
DB 1381 gagcctgtatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaaatgaa 1423

RESULT 15

BOVME

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

FEATURES

Source

Location/Qualifiers

1. 1417

/organism="Bos taurus"

/db_xref="taxon:9913"

/cell_type="Secretory cell"

/dev_stage="25 weeks"

/tissue_type="Tongue serous gland"

73. 129

73. 1266

/EC_number="3.1.1.3"

/product="pregastric esterase"

/protein_id="AAB57037.1"

/db_xref="GI:600757"

/translation="MWWLLVTVCIHNSGNAFCFLGIKRNPEASNMVNSQMSISWYGP

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2002, 21:08:53 ; Search time 166.02 Seconds

(without alignments)
6553.100 Million cell updates/sec

Title: US-09-578-063-46

Perfect score: 1269

Sequence: 1 atgttggaacctgtcgaag.....gacggtgtgagccgcatgtt 1269

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues 1861242

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_1101.*

1:	/SID58/gcgdata/geneseq/NA1980.DAT.*
2:	/SID58/gcgdata/geneseq/NA1981.DAT.*
3:	/SID58/gcgdata/geneseq/NA1982.DAT.*
4:	/SID58/gcgdata/geneseq/NA1983.DAT.*
5:	/SID58/gcgdata/geneseq/NA1984.DAT.*
6:	/SID58/gcgdata/geneseq/NA1985.DAT.*
7:	/SID58/gcgdata/geneseq/NA1986.DAT.*
8:	/SID58/gcgdata/geneseq/NA1987.DAT.*
9:	/SID58/gcgdata/geneseq/NA1988.DAT.*
10:	/SID58/gcgdata/geneseq/NA1989.DAT.*
11:	/SID58/gcgdata/geneseq/NA1990.DAT.*
12:	/SID58/gcgdata/geneseq/NA1991.DAT.*
13:	/SID58/gcgdata/geneseq/NA1992.DAT.*
14:	/SID58/gcgdata/geneseq/NA1993.DAT.*
15:	/SID58/gcgdata/geneseq/NA1994.DAT.*
16:	/SID58/gcgdata/geneseq/NA1995.DAT.*
17:	/SID58/gcgdata/geneseq/NA1996.DAT.*
18:	/SID58/gcgdata/geneseq/NA1997.DAT.*
19:	/SID58/gcgdata/geneseq/NA1998.DAT.*
20:	/SID58/gcgdata/geneseq/NA1999.DAT.*
21:	/SID58/gcgdata/geneseq/NA2000.DAT.*
22:	/SID58/gcgdata/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1269	100.0	1269	22	Human TANGO 294 CP
2	1269	100.0	22	AAE45132	Human TANGO 294 CP
3	526.2	41.5	245	22	AAH29J06
4	489.2	38.6	1378	14	AAO42310
5	470	37.0	1367	7	AAH60685
6	470	37.0	1367	7	AAH60566
7	470	37.0	1367	17	AAH58916
8	468.4	36.9	1695	22	AAH57420
9	464.4	36.6	1336	6	AAH50385
10	444.2	35.0	1528	15	AAO68388
11	444.2	35.0	1528	17	AAH58915

12	444.2	35.0	1531	15	AAO68389	Canine gastric lip
13	286	22.5	699	22	AAF28679	Human protein HP03
14	286	22.5	1308	22	AAF28689	Human protein HP03
15	257.8	20.3	617	21	AAA44349	Human secreted exp
16	193.2	15.2	502	21	AAA43744	Mouse secreted exp
17	100.2	7.9	402	22	AAH57212	Human stomach spec
18	100	7.9	1718	21	AAZ34958	Soybean acid triac
19	89.4	7.0	1483	21	AAZ34956	Rice acid triacylg
20	84.8	6.7	859	21	AAZ34953	Corn acid triacylg
21	72	5.7	1438	21	AAZ34959	Soybean acid triac
22	68	5.4	936	22	AAH58252	Oligonucleotide D1
23	68	5.4	936	22	AAH58254	Oligonucleotide D1
24	68	5.4	936	22	AAH58257	Oligonucleotide D1
25	68	5.4	936	22	AAH58259	Oligonucleotide D2
26	68	5.4	936	22	AAH58262	Oligonucleotide D2
27	68	5.4	936	22	AAH58255	Oligonucleotide D1
28	66.2	5.2	936	22	AAH58252	Oligonucleotide D1
29	66.2	5.2	936	22	AAH58254	Oligonucleotide D1
30	66.2	5.2	936	22	AAH58257	Oligonucleotide D1
31	66.2	5.2	936	22	AAH58259	Oligonucleotide D2
32	66.2	5.2	936	22	AAH58262	Oligonucleotide D2
33	66.2	5.2	936	22	AAH58255	Oligonucleotide D1
34	55.8	4.4	705	21	AAZ34952	Catalpa acid triac
35	54	4.3	549	21	AAH3827	Cat flea hindgut a
36	51	4.0	509	21	AAZ34954	Corn acid triacylg
37	48.2	3.8	178	19	AAV35431	Cell death regulat
38	46.4	3.7	43676	21	AAZ60606	Nucleotide sequenc
39	38.6	3.0	1664976	19	AAV21209	Methanococcus jam
40	36.8	2.9	13254	21	AAA32165	Complete nucleotid
41	36.6	2.9	244	22	AAH58238	Oligonucleotide D1
42	36.6	2.9	3271	19	AAV00060	Human G-protein co
43	36.6	2.9	3282	17	AAH40914	DNA encoding VH re
44	36.6	2.9	3282	21	AAA31024	Modified heavy cha
45	36.6	2.9	3282	21	AAA31038	B12 IgG1 preferred

ALIGNMENTS

RESULT 1					
ID	AAH45132	standard: cDNA: 1269 BP.			
XX	AAH45132;				
XX	30-MAR-2001 (first entry)				
DE	Human TANGO 294 ORF.				
XX					
KW	Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;				
KW	central nervous system; focal brain disorder; bipolar affective disorder;				
KW	global diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;				
KW	senile dementia; Huntington's disease; amyotrophic lateral sclerosis;				
KW	Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;				
KW	neuropsychiatric; psychoactive substance use; anxiety; ss.				
XX					
OS	Homo sapiens.				
PN	WO200077239-A2.				
XX					
PD	21-DEC-2000.				
XX					
XX	4-MAY-2000; 2000WO-US14858.				
PP	14-JUN-1999; 99US-0333159.				
PR					
XX	(MILL-) MILLENNIUM PHARM INC.				
PA					
XX	McCarthy SA, Fraser CC, Sharp JD, Barnes TM;				
PI	WPI: 2001-032313/04.				
DR	P-PSDB: AAB66065.				
DR					
XX					

PT TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
PT screening assays and diagnostic assays and for the treatment of
PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
PT disease.

Claim 1; Fig 6; 359pp; English.

CC The present invention relates to TANGO or INTERCEPT proteins and coding
CC sequences (see AAF45121-F45136 and AAF45138-F45139 and AAB6031-B66057,
CC AAB6064-B66083 and AAB6085). The TANGO/INTERCEPT proteins and coding
CC sequences are useful for the treatment of neurological disorders such as
CC central nervous system (CNS) disorders, CNS-related disorders, focal
CC brain disorders, global-diffuse cerebral disorders and other
CC neurological and cerebrovascular disorders. The CNS disorders include
CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
CC autonomic function disorders such as hypertension and sleep disorders,
CC neuropsychiatric disorders, psychoactive substance use disorders,
CC anxiety, and bipolar affective disorder.

XX Sequence 1269 BP; 358 A; 268 C; 309 G; 334 T; 0 other:

Query Match 100.0%; Score 1269; DB 22; Length 1269;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgttggaacctgttcagaacagtgatgtgtccacacagaatgtaattgtgcttcg 60
Db 1 atgttggaacctgttcagaacagtgatgtgtccacacagaatgtaattgtgcttcg 60
QY 61 attctgggtggtatatttccagaagaatgtgaattcagtaacatagccaataagct 120
Db 61 attctgggtggtatatttccagaagaatgtgaattcagtaacatagccaataagct 120
QY 121 gtggaccagagaacatcatatattagtgaaatcatccacatcaagggtatcccggt 180
Db 121 gtggaccagagaacatcatatattagtgaaatcatccacatcaagggtatcccggt 180
QY 181 gaggaataatgaagtcgcgaactgaagatggtatataccttctgttaacaggaattccctga 240
Db 181 gaggaataatgaagtcgcgaactgaagatggtatataccttctgttaacaggaattccctga 240
QY 241 ggctatgtgcaactaagaagaacaggtcgcagcctgtgtgtactgacagatgacctga 300
Db 241 ggctatgtgcaactaagaagaacaggtcgcagcctgtgtgtactgacagatgacctga 300
QY 301 gttagaggtgtacgaactgattccaactgcccacaataagcctgggcttattctg 360
Db 301 gttagaggtgtacgaactgattccaactgcccacaataagcctgggcttattctg 360
QY 361 gcaaatgctggttttgacgtgtgagtggagacagcagggaacgcctgtgtccgaaaa 420
Db 361 gcaaatgctggttttgacgtgtgagtggagacagcagggaacgcctgtgtccgaaaa 420
QY 421 cacaagaacacttccatagaccaaagatgagttcctggccttcagttatgagatgagct 480
Db 421 cacaagaacacttccatagaccaaagatgagttcctggccttcagttatgagatgagct 480
QY 481 aggtttgaacctctcagtgatataaacttttttgcagaaaaaaggccaggaataaatc 540
Db 481 aggtttgaacctctcagtgatataaacttttttgcagaaaaaaggccaggaataaatc 540
QY 541 tattaatgtcggtatatacagaggcaccacatgggtcttattgcatlccaccatgcca 600
Db 541 tattaatgtcggtatatacagaggcaccacatgggtcttattgcatlccaccatgcca 600
QY 601 gagctggtcagaataatcaaatgtatttgccttagcaccatagccactgttaagcat 660
Db 601 gagctggtcagaataatcaaatgtatttgccttagcaccatagccactgttaagcat 660
QY 661 gcaaaaagcccgagacaaattttgtctgcagatatgagatcaaggagatttt 720
Db 661 gcaaaaagcccgagacaaattttgtctgcagatatgagatcaaggagatttt 720

Db 661 gcaaaaagcccgagacaaattttgtctgcagatatgagatcaaggagatttt 720
QY 721 ggcacaaagaattctgtatcagacagatcttcagacaaattgttattacccttgc 780
Db 721 ggcacaaagaattctgtatcagacagatcttcagacaaattgttattacccttgc 780
QY 781 ggcagagtgattcttgcagagatttgaatataatcatgttaacttcgggtgattaac 840
Db 781 ggcagagtgattcttgcagagatttgaatataatcatgttaacttcgggtgattaac 840
QY 841 accacacatagcaacagagccggcagagtgatatactccacacatctgtcggagaca 900
Db 841 accacacatagcaacagagccggcagagtgatatactccacacatctgtcggagaca 900
QY 901 tctgtgcaaatatcttaccactgtgagccagcagtgaaattctgtgtaactccggcatt 960
Db 901 tctgtgcaaatatcttaccactgtgagccagcagtgaaattctgtgtaactccggcatt 960
QY 961 gactgggggtgtgagacacaaatcttgaaaaaatgcaatcagccaactcctgttaagtac 1020
Db 961 gactgggggtgtgagacacaaatcttgaaaaaatgcaatcagccaactcctgttaagtac 1020
QY 1021 agagtcagagatatgacggtccctacagcaatgtggaacagaggtcagagctgttca 1080
Db 1021 agagtcagagatatgacggtccctacagcaatgtggaacagaggtcagagctgttca 1080
QY 1081 aatccagaagaacgtgaaatgctgtctctgtgaggtgacacacatctaccataagaat 1140
Db 1081 aatccagaagaacgtgaaatgctgtctctgtgaggtgacacacatctaccataagaat 1140
QY 1141 attcctgaatgtggtcagcgtggaattcatctgtgggtttgagatgctctaccagatgac 1200
Db 1141 attcctgaatgtggtcagcgtggaattcatctgtgggtttgagatgctctaccagatgac 1200
QY 1201 aatgaatcatcatctcgtgacagcagagagagacacacattccacggagcgtgtgag 1260
Db 1201 aatgaatcatcatctcgtgacagcagagagagagacacacattccacggagcgtgtgag 1260
QY 1261 gccgtattg 1269
Db 1261 gccgtattg 1269

RESULT 2
AAF45131
ID AAF45131 standard; cDNA; 2044 BP.
XX
AC AAF45131;
DT 30-MAR-2001 (first entry)
XX
DE Human TANGO 294 cDNA.
XX
KW Gene therapy; TANGO protein; INTERCEPT protein; neurological disorder;
KW central nervous system; focal brain disorder; bipolar affective disorder;
KW global-diffuse cerebral disorder; cerebrovascular; Alzheimer's disease;
KW senile dementia; Huntington's disease; amyotrophic lateral sclerosis;
KW Parkinson's; Gilles de la Tourette's syndrome; hypertension; sleep;
KW neuropsychiatric; psychoactive substance use; anxiety; ss.
OS Homo sapiens.
XX
PN WO200077239-A2.
PD 21-DEC-2000.
XX
PF 24-MAY-2000; 2000WO-US14858.
XX
PR 14-JUN-1999; 99US-0333159.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI McCarthy SA, Fraser CC, Sharp JD, Barnes TM;

XX WPI: 2001-032313/04.
 DR P-PSDB: AAB66065.
 DR
 XX TANGO and INTERCEPT nucleic acids, proteins, and antibodies, useful for
 PT screening assays and diagnostic assays and for the treatment of
 PT neurological diseases such as Alzheimer's, Parkinson's and Huntington's
 PT disease -
 XX
 XX Claim 1; Fig 6; 359pp; English.
 PS
 CC The present invention relates to TANGO or INTERCEPT proteins and coding
 CC sequences (see AAF5121-F45136 and AAF5138-F45139 and AAB66031-B66057,
 CC AAB66064-B66083 and AAB66085). The TANGO/INTERCEPT proteins and coding
 CC sequences are useful for the treatment of neurological disorders such as
 CC central nervous system (CNS) disorders, CNS-related disorders, focal
 CC brain disorders, global-diffuse cerebral disorders and other
 CC neurological and cerebrovascular disorders. The CNS disorders include
 CC Alzheimer's disease, senile dementia, Huntington's disease, amyotrophic
 CC lateral sclerosis, Parkinson's, Gilles de la Tourette's syndrome,
 CC autonomic function disorders such as hypertension and sleep disorders,
 CC neuropsychiatric disorders, psychoactive substance use disorders,
 CC anxiety, and bipolar affective disorder.
 CC
 XX
 SO Sequence 2044 BP; 618 A; 401 C; 460 G; 565 T; 0 other:
 Query Match 100.0%; Score 1269; DB 22; Length 2044;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atgttggaacctgttcaagaacatgtgattgtctcacacaagaatggaatgtgcttctg 60
 Db 126 atgttggaacctgttcaagaacatgtgattgtctcacacaagaatggaatgtgcttctg 185
 QY 61 attctgtgagcgtatatttccagaagaatgtgaattcagatcatatgccaactaaagct 120
 Db 186 attctgtgagcgtatatttccagaagaatgtgaattcagatcatatgccaactaaagct 245
 QY 121 gtggaaccagaacatcatgaataattagtaaatcatccaacatcaagctatccctgt 180
 Db 246 gtggaaccagaacatcatgaataattagtaaatcatccaacatcaagctatccctgt 305
 QY 181 gaggaattgaagtcgcgaactgaagatgtgatatccttctgttaaaagattccttga 240
 Db 306 gaggaattgaagtcgcgaactgaagatgtgatatccttctgttaaaagattccttga 365
 QY 241 ggctagtgcacaactaaagaacagaggtccaggcctgtgtgttactgacagatggccta 300
 Db 366 ggctagtgcacaactaaagaacagaggtccaggcctgtgtgttactgacagatggccta 425
 QY 301 gtggaagtgctgaactgtgattccaactgtcccaacaatagcctggcttcatctg 360
 Db 426 gtggaagtgctgaactgtgattccaactgtcccaacaatagcctggcttcatctg 485
 QY 361 gcaaatgtgttttgaagtgttgatgtagggaacagggaaagcctgtgtctgaaaa 420
 Db 486 gcaaatgtgttttgaagtgttgatgtagggaacagggaaagcctgtgtctgaaaa 545
 QY 421 cacaagaacctccatagacacaaagatagttctgggcttccagttaagatgagct 480
 Db 546 cacaagaacctccatagacacaaagatagttctgggcttccagttaagatgagct 605
 QY 481 aggttgaccttctgcagtgataaacttattttgcagaaaacggccaggaagaagtc 540
 Db 606 aggttgaccttctgcagtgataaacttattttgcagaaaacggccaggaagaagtc 665
 QY 541 tatatgctggcctttacacagggacacacacaggttattgatttccacacatgca 600
 Db 666 tatatgctggcctttacacagggacacacacaggttattgatttccacacatgca 725
 QY 601 gactgtgctcagaataaattatttctttagaccacatagcactgttaagat 660
 Db 726 gactgtgctcagaataaattatttctttagaccacatagcactgttaagat 785

Db 726 gactgtgctcagaataaattatttctttagaccacatagcactgttaagat 785
 QY 661 gcaaaaagcccgagaccaaaattttgtgtctgccagatcatgatcaaggatgtt 720
 Db 786 gcaaaaagcccgagaccaaaattttgtgtctgccagatcatgatcaaggatgtt 845
 QY 721 ggcataaagaattttgtatcacagccagatcttccagaacattgttatttacttgt 780
 Db 846 ggcataaagaattttgtatcacagccagatcttccagaacattgttatttacttgt 905
 QY 781 ggcagtgattcttgatcacagatttgtagtaatacatgtaactctgggtgattcaac 840
 Db 906 ggcagtgattcttgatcacagatttgtagtaatacatgtaactctgggtgattcaac 965
 QY 841 accaacaatgatgaatgagccgagcaagtgtatatgtctgccacactctgtcggaca 900
 Db 966 accaacaatgatgaatgagccgagcaagtgtatatgtctgccacactctgtcggaca 1025
 QY 901 tctgtgcaaatattcttactgagccagggcagtgaaattctgttgaactccggcatt 960
 Db 1026 tctgtgcaaatattcttactgagccagggcagtgaaattctgttgaactccggcatt 1085
 QY 961 gactgggagtgagaccacaataatctggaataatgcaatccgaactcctgttaagttac 1020
 Db 1086 gactgggagtgagaccacaataatctggaataatgcaatccgaactcctgttaagttac 1145
 QY 1021 agagtcagagatgatgacggctcccaagaacatgtggacaggaagtcaggactgttca 1080
 Db 1146 agagtcagagatgatgacggctcccaagaacatgtggacaggaagtcaggactgttca 1205
 QY 1081 aatccagaagaacgrrgaataatctgtcctctgagtgaccacacatctactacaagaat 1140
 Db 1206 aatccagaagaacgrrgaataatctgtcctctgagtgaccacacatctactacaagaat 1265
 QY 1141 attcctgaatggctcacgtgatttcatctcggggtttagatgtctccacaggtatgtac 1200
 Db 1266 attcctgaatggctcacgtgatttcatctcggggtttagatgtctccacaggtatgtac 1325
 QY 1201 aatgaatcatcatctgtatgtcagcaggaagagacacacatttccaggagcgtgtgag 1260
 Db 1326 aatgaatcatcatctgtatgtcagcaggaagagagacacacatttccaggagcgtgtgag 1385
 QY 1261 gccgtattg 1269
 Db 1386 gccgtattg 1394

RESULT 3
 AAH02906
 ID AAH02906 standard; DNA; 2493 BP.
 XX
 XX AAH02906;
 AC
 XX
 XX 15-JUN-2001 (first entry)
 DT
 XX
 DE Human shear stress-response coding sequence SEQ ID NO: 65.
 KW Human; shear stress-response protein; vascular disease;
 OS arteriosclerosis; ds.
 XX
 XX Homo sapiens.
 XX
 XX WO200125427-A1.
 XX
 XX 12-APR-2001.
 PD
 XX
 XX 02-OCT-2000; 2000WO-JP06840.
 PF
 XX
 XX 01-OCT-1999; 99JP-0280976.
 PR
 XX
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA
 PA (NOJIMA) NOJIMA H.
 XX

PT treating conditions linked to gastric lipase deficiency, such as
 PT mucoviscidiosis and pancreatic exocrine insufficiency
 PS Claim 1; Fig 6; 31pp; French.

XX mRNA was isolated from rabbit gastric mucosal cells, converted to
 CC cDNA and fragments cloned in pUC18. Recombinant plasmids were used
 CC to transform E.coli MM294 and cells screened for reactivity with
 CC probes corresp. to known parts of the RGL sequence. One positive
 CC clone cont. plasmid pJ0101 which contd. a 1.35 kb psiI sequence.
 CC This fragment was isolated, subjected to PCR and the amplification
 CC prod. ligated with the 2.2 kb BglII-SacI fragment of pR276 (contg.
 CC a synthetic Tac promoter and transcription stop signal). The
 CC recombinant plasmid was designated pRGLN2.1.

XX Sequence 1378 BP; 397 A; 287 C; 276 G; 418 T; 0 other;

Query Match 38.6%; Score 489.2; DB 14; Length 1378;
 Best Local Similarity 65.0%; Pred. NO.3.7e-145;
 Matches 722; Conservative 0; Mismatches 388; Indels 0; Gaps 0;

QY 125 acccagaagcatcatgaatattagtaaatccacacccaagctatccctgtgaag 184
 DB 136 accctgaagtgaaatgaatattagtcagatgtattctctgtggataccagaatgaa 195
 QY 185 aatatgaagtcgaactgaagatgggtatatccttctgttaacagaattcctcgagcc 244
 DB 196 aatatgaagtcgtactgaagatggtatatccttctgtgttaacagaattccttatgga 255
 QY 245 tatgtaaaccttaagagacaggtctccaggtgtgtgtgtgtgtgtgtgtgtgtgtgt 304
 DB 256 agaaaaattcaaggagacagagccagaccgcgtcattcttgcagcaggtgtgtgtgt 315
 QY 305 gaagtcctagacactgattccacactgcacaaatagccttgggtattctctgcag 364
 DB 316 catcagcctcaaaactgtatctcactgcacaaacagcctgtcttattctgtcag 375
 QY 365 atgtgtgttttgacgt 424
 DB 376 atgt 435
 QY 425 agacactctccatagacagaatgaggtctgtgtgtgtgtgtgtgtgtgtgtgtgtgt 484
 DB 436 tatacatctccacagacagctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 495
 QY 485 ttgacctctctgagtgatataacttattttgcagaaaacgggcagaaagatctatt 544
 DB 496 atgaccttcagcagacattgacttctgttaagaaatgtgcagaaagatctgact 555
 QY 545 atgtgtgtctatcagagcagacaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 604
 DB 556 atgtgtgtctatcagagcagacaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 615
 QY 605 ttgctgaagaaatcaaatgtatttttgcagccacatgcagcagctgttaagatgaa 664
 DB 616 tggctgaagaaatcaaatgtatttttgcagccacatgcagcagctgttaagatgaa 675
 QY 665 aaagcccgagacaaattttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 724
 DB 676 aaagcccgagacaaattttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 735
 QY 725 aaagaattctgtatcagacagattctcagacaattgtatttatttatttatttggcc 784
 DB 736 acaaaatattctacccacaaattcttctgtatcaattcttgcacacaaatgtgtgtcc 795
 QY 785 aggtgtgtctgt 844
 DB 796 gtgagcaactgaatgttaatttgcagcaatgcttatttatttatttatttatttggcc 855
 QY 845 acatatgaacatgagccgagcagatgtatattgtgtgtgtgtgtgtgtgtgtgtgtgtgt 904
 DB 856 caaactgtacacatgagcgt 915

QY 905 tgcnaaatattctacactgagccagcagtgatattctgtgtgaactccgggacttgact 964
 DB 916 tccaacaacatgctgcagcagccagcagctgttaattctggaatttccaagctttaat 975
 QY 965 ggggagtgagacaaaatcttgaaaatgcataccagcaactctgttaagtagaag 1024
 DB 976 ggggagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1035
 QY 1025 tcaagatattgagcgtctccacagcaatgtgacagaggttcagagctgtcttaac 1084
 DB 1036 tggccgcagatgagtgccaatgtgcaggtgtgagaggtgtgagcagctgtgtgtgt 1095
 QY 1085 cagaagcagcgaatgagcgtctctcagagtgccacacatctcactcaagaatattc 1144
 DB 1096 cccaagatgttgaccttgccttcccaaatcttcaatcattccacaaagagatc 1155
 QY 1145 ctgaatggctcagcagtgatctcagcaggtgtgtgtgtgtgtgtgtgtgtgtgtgt 1204
 DB 1156 ttcataatcactgt 1215
 QY 1205 aatcatcactgtatgcagcagagaga 1234
 DB 1216 aaattattctatgatgtgcaaaagataaa 1245

RESULT 5
 AAN60685
 ID AAN60685 standard; DNA: 1367 BP.

AC AAN60685;

DF 08-JUN-1991 (first entry)

XX Sequence encoding pregastric lipase.

KW Precursor polypeptide; secretion vector; enzyme; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT sig_peptide 47..103

FT mat_peptide 104..1243

FT /tag- b

XX MO8603778-A.

XX PD 03-JUL-1986.

XX PF 23-DEC-1985; 85MO-GB00599.

XX PR 21-DEC-1984; 84GB-0032483.

XX PR 23-DEC-1985; 85MO-GB00599.

XX PR 01-JAN-1986; 86GB-0019568.

XX PA (BREN-) BREWING RES FOUND.

XX PA (TUBB/) TUBB R S.

XX PI Tubb RS;

XX DR WPI, 1986-182910/28.

XX DR P-PSDB; AAP60724.

XX PT New precursor polypeptide of defined sequence - and corresp. DNA

XX used to transform hosts for prodn. of the polypeptide

XX Example; Fig 5; 60pp; English.

XX The inventors claim a DNA sequence encoding AAP60723 linked to a

CC promoter upstream and a gene for a polypeptide downstream.

CC Particular examples are the yeast enzyme AMG, the mammalian enzyme,

CC gastric lipase and the mammalian lymphokine, interferon-alpha2.

XX Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;
SQ

Query Match 37.0%; Score 470; DB 7; Length 1367;
Best Local Similarity 64.2%; Pred. No. 4,8e-139;
Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;

QY 119 ctgtgaccagagatcatcatgatatatgtaaatcaatcaacatcaaggctaccc 178
DB 123 ctggaagccctgagtgactatgacattagtcagatgattcttctgtaacagatcc 182
QY 179 gtgaggaataatgagtcgcaactgcaagatggtatccctctctgtaacagatcc 238
DB 183 atgaaagaatatgagtgactgagtcgcaagatggtatctctgtaacagatcc 242
QY 239 gaggcctagtgcacctaagaagacaggtccagcctgtgtgttactgacatgccc 298
DB 243 atggaagaagaataatcagaggaatcacagcagacacctgtgtgttcttgcagatggt 302
QY 299 tagttgaggtgtgtagcaactgattccaaactgcccacaataagcctgggtctatc 358
DB 303 tctctgcacagccacaactgattccaaactgcccacaacacgacctgtgcttctc 362
QY 359 tggcagatgctggttttgacgtgtgtgagtgagcagcagggaggaacgacctgtgcga 418
DB 363 tggcagatgctggttttgacgtgtgtgagtgagcagcagggaggaacacacctgtgcga 422
QY 419 aacacaagaacactcctccatagaccagaatgagttctgggcttcagtatgagatgg 478
DB 423 gaactctgtactatctaccagatctgaaattcctgggtccttcagcttgcagatgag 482
QY 479 ctgagttgaacctctgtagtgatgataaacttttttgcagaanaacggccagaaaga 538
DB 483 ctaaaatagacctccagacacaactgacttctgtaagaaactgacagaagcagc 542
QY 539 tctatatgtcgcgtatctacacagcagccacacatggtcttattgtcatttccacatgc 598
DB 543 taacatattgtgacatctccacagcagccacacatggtcttattgtcatttccacatgc 602
QY 599 cagaagctggtcagaaatcaaaatgatttcttgcacccaatagccactgttaagc 658
DB 603 ccagcctgtgctaaagaatcaaaactctatgctctcgtctgctcagctgtgaagt 662
QY 659 atgcaaaaagcccgagcaaaaattttgtgtgctgcagatagatagatcaagagttgt 718
DB 663 atcaaaaagccctataaacaacttagattgttcccaatccctcccaagttatata 722
QY 719 ttgacaaaagaagaatttctgtatcagacagatcttcagacaaactgttattacctt 778
DB 723 ttgtgtgacaaaatattctacccacacacttcttgaatcaattcttctgctgaagtg 782
QY 779 gtggcaggtgattctctgtacagattgtgataatcatcattgtaactctgggtgattca 838
DB 783 gctcccgtagatgtcgaattcttcttgcagcatgcttattataatttggattgtg 842
QY 839 acacacaaatctgaacatgagccgagcaggtatatactgcccacactctgtcggaga 898
DB 843 acagtaagaacttaacacagagtcgtgtgattatctatacatcaatcaacagagaa 902
QY 899 catctgtgcaaaaatctctacactgagccagcaggtgaattctgtgtgactccgggac 958
DB 903 ctctctgtcaaaaatctctacactgagccagcaggtgttaagctgtgaaatccagact 962
QY 959 ttgactgggggaggtgagaccaaaatctggaataatgcaatcaacactcctgttaagt 1018
DB 963 atgactgggggaggtgagaccaaaatctggaataatgcaatcaacactcctgttaagt 1022
QY 1019 acagagctgagagatagaggtccctacagcagatggtgagcagaggttcagagctctt 1078
DB 1023 acaaatgtgacacacatgaaatgtacaaatgtgagcagtggtgagcagagctgtg 1082
QY 1079 caaatcagaagaagctgaaatgtctctctgaggtgagcacaacctcactacataaga 1138

DB 1083 ctgaccocccaagatgtgtgctcttcttcccaaaactccccaactctatttaccacaag 1142
QY 1139 atattcctgaatgtgctcagctgagttcatctcgtgggttggatgctctccacgtaigt 1198
DB 1143 agattcctctttacacacttgactttatctcgtggcaatgagtgccctcaagaagttt 1202
QY 1199 acaatgaatcatccatcgtat 1220
DB 1203 acaatgacattgttctatgat 1224

RESULT 6

AA060566
ID AA060566 standard; DNA; 1367 BP.

AA060566;
22-AUG-1991 (first entry)

Sequence encoding human pregastric lipase.

Cystic fibrosis therapy; enzyme; lipase deficiency; ss.

Homo sapiens.

Key

CDS Location/Qualifiers

FT 47..103

FT mat_peptide 104..1243

FT /*tag= b

PN W08601532-A.

PD 13-MAR-1986.

PF 15-AUG-1985; 85WO-GB00364.

PR 21-AUG-1984; 84GB-0021210.

PR 15-AUG-1985; 85WO-GB00364.

PR 01-JAN-1986; 86GB-000897.

PA (CELL-) CELTECH LTD.

PI (LOWE/) LOWE P A.

PI Lowe PA.

DR WPI; 1986-081634/12.

DR P-PSDB; AAP60658.

PT New gastric lipase protein, esp. of human origin - for treating

PT lipase deficiency, and DNA sequences coding for it

PS Disclosure; Fig 3; 39p; English.

PS The inventors claim a pregastric lipase protein and a gene encoding

CC it. Gastric lipase protein is useful for oral administration to

CC treat lipase deficiency, e.g. cystic fibrosis or pancreatitis.

CC Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;

SQ

Query Match 37.0%; Score 470; DB 7; Length 1367;
Best Local Similarity 64.2%; Pred. No. 4,8e-139;
Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;

QY 119 ctgtgaccagagatcatcatgatatatgtaaatcaatcaacatcaaggctaccc 178
DB 123 ctggaagccctgagtgactatgacattagtcagatgattcttctgtaacagatcc 182
QY 179 gtgaggaataatgagtcgcaactgcaagatggtatccctctctgtaacagatcc 238
DB 183 atgaaagaatatgagtgactgagtcgcaagatggtatctctgtaacagatcc 242

QY 239 gagcgctatgtcaactaagaagacaggttcacagcgctgtgtgttactgcagcatgccc 298
 DB 243 atggaagaagaataatcaagggaatacagagccgtgtgttcttgcagcatggtt 302
 QY 299 tagttggaggtgctagcaactggatttccaaacctgcacacatgagctggtgcttacc 358
 DB 303 tgccttgatcgcagcaaacatggtatttccaaacctgcagcaaacagctgttccatc 362
 QY 359 tggcagatgctggttttgcagctgtgtgaggaagacagaggggaaagcctgtgtcga 418
 DB 363 tggcagatgctggttttgcagctgtgtgaggaagacagaggggaaagcctgtgtcga 422
 QY 419 aacccaagaacctctcccttagaccagaatggttctggtcttccgtatgtagatg 478
 DB 423 gaacattgtactatcaccaagattcagttgatttctggtcttccgtatgtagatg 482
 QY 479 ctaggtttgacctctctcagtgatgaacttatttgcagaaacggcgaggaaga 538
 DB 483 ctaaatatgacctctccagccacatcgactcattgttaagaagaacttgacagaagc 542
 QY 539 tctatattgctcggtattcacagggcacacatggtgcttatttgcatttccacatgc 598
 DB 543 tactactatgttggccattccagggcacacatggttatttcttccatccacatc 602
 QY 599 cagagctgtgctcagaanaatcaaatgtatttgccttagcaccatagccactgttaagc 658
 DB 603 ccagcctgtctaaagaatacaaaccttctatgctctagctcctgttgcactgtgaag 662
 QY 659 atgcataaagcccccgggaccaaatttctgtgtcgcagatgtgtatcaaggatgt 718
 DB 663 atacaaaagcccttaaacaaacttagattgttcttccaaacctctcctcaagttat 722
 QY 719 ttggcaaaaagaattcttctatagaccagattctcagaacattgttatttccctt 778
 DB 723 ttgttgacaataatatttccacacacacattcttgcataatttcttgcattgaggt 782
 QY 779 gtggcaggtgatttcttgcagatttgcagatttgcagatttgcagatttgcagatt 838
 DB 783 gctcccgagatgtgatttcttgcagatttgcagatttgcagatttgcagatttgcag 842
 QY 839 acacacaataatgacatgagccgagcaagtgtatgtatgtgtccacacttctgtgaa 898
 DB 843 acagtaagaacttaccacagagctgtgagtgatctatcacaataatccagcagaa 902
 QY 899 catcgtgtcaaaatatttacccttgagcagcaggtgaattctgtgacccgggcat 958
 DB 903 ctctgttcaaaaatatttcccttgagcagcaggtgaattctgtgacccgggcat 962
 QY 959 ttgactggggaggtgagcaaaaatctcgaaataatgcatacagccaactcctgtaa 1018
 DB 963 atgactggggaggtgagcaaaaatctcgaaataatgcatacagccaactcctgtaa 1022
 QY 1019 acagatcagagatgacggttccctacagcaaatgtgacagagaggtccagagctt 1078
 DB 1023 acaatgtacagcattgatttaccatgtcagttgaaagcgttgcaagagccgttgg 1082
 QY 1079 caaatcagagagcgtgaaatgctgctcctcagaggtgacacactcattccataa 1138
 DB 1083 ctgacccccaagatgtgcttcttgccttccaaactccccaacttatttccccaag 1142
 QY 1139 atattcctgaatgggtcagtgatcattcattcagtggttgcctccatccatg 1198
 DB 1143 agattccttttaacatacacttgagcttattcctgagcaatgagcttccatccat 1202
 QY 1199 acaatgaataatcatctgtat 1220
 DB 1203 acaatgacattgttctatgat 1224

RESULT 7
 AAT58916
 ID AAT58916 standard; DNA; 1367 BP.

XX AC AAT58916;
 XX DT 19-AUG-1997 (first entry)
 XX DE Human gastric lipase coding sequence.
 XX KW Duodenal; gastric; lipase; transgenic; plant; recombinant; extract; ss;
 KW food; absorption; fat; pancreatic; cystic fibrosis; exocrine; dairy;
 KW hydrolysis; trans-esterification; substrate; enzyme; biofuel; human.
 OS Homo sapiens.
 XX FH Key location/Qualifiers
 XX FT CDS 47..1243
 XX FT /tag= a
 XX FT /product= gastric lipase
 XX FT sig_peptide 47..103
 XX FT /tag= b
 XX FT mat_peptide 104..1240
 XX FT /tag= c
 XX PN W09633277-A2.
 XX PD 24-OCT-1996.
 XX PE 19-APR-1996; 96WO-FR00606.
 XX PR 20-APR-1995; 95FR-0004754.
 XX PA (BIOC-) BIOCEM SA.
 XX PA (LJOU) INST RECH JOUVEINAL.
 XX PI Baudino S, Benicourt C, Cudrey C, Gruber V, Lenee P;
 XX PI Merot B;
 XX DR WPI: 1996-485783/48.
 XX DR P-PSDB: AAM09383.
 XX PT Recombinant human or canine pre-duodenal lipase proth. in transgenic
 XX PT plants - useful for facilitating absorption of fat, as bio-catalysts
 XX PT and for prodn. of bio-fuel
 XX PS Claim 3; Fig 4; 130pp; French.
 XX CC This is the nucleotide sequence encoding the human pre-duodenal (i.e.
 CC gastric) lipase enzyme. The sequence can be used to generate transgenic
 CC plants producing recombinant lipase in an enzymatically active form.
 CC CC Alternatively bases encoding amino acids 20-23 or 20-73 (1-4 or 1-54 of
 CC the mature protein, respectively) can be deleted to form the derivatives
 CC designated delta-4 or delta-54 respectively. Plants, or their extracts,
 CC expressing the lipases or the truncated derivatives, can be used:
 CC (a) as pharmaceuticals or food to facilitate absorption of fat, either
 CC in healthy subjects or in patients with inadequate levels of
 CC gastric/pancreatic lipase e.g. those with cystic fibrosis or exocrine
 CC pancreatic insufficiency. In the elderly or in patients undergoing medical
 CC treatment which alters fat adsorption; (b) for performing industrial or
 CC agricultural reactions, e.g. in processing of fats or in the dairy
 CC industry, for hydrolysis or transesterification reactions, etc., where
 CC the plant material may provide both enzyme and substrate. The transgenic
 CC plants can also be used for biofuel production.
 XX SQ Sequence 1367 BP; 397 A; 284 C; 269 G; 417 T; 0 other;

Query Match 37.0%; SCORE 470; DB 17; Length 1367;
 Best Local Similarity 64.2%; Pred. No. 4.8e-139;
 Matches 707; Conservative 0; Mismatches 395; Indels 0; Gaps 0;

QY 119 ctgtgaccagaagcattcatgaatattagtgaaatcatcaacataagctatccct 178
 DB 123 ctggaagccctgaagtgactatgaacattatgacatgattactatctggggataccaa 182

OY	179	gtgaggaaatagaagctgcgaacatgaagaatgtgatactaccttctggttaacagaagattctc	238
Db	163	atgaagaataatgaagatgtgtgaactgaatgaatgtgtatattcttctgaagataaatagaattcttc	2422
OY	229	gaggctctagtcgaacctgaagaagaacaggtttccagagctctgtgtgttactctgaacaatgccc	2368
Db	243	atgggaagaagaataattccagagatatacagccagagactcgtgtgttttttgcagatggtt	3020
OY	299	taagtgtgaggtgtctagaacacgtggtttccaaacctgcggccaacataagctgtgcttaattc	358
Db	303	tgcttgctgcacgccaacaacacgtggtttccaaacctgcggccaacaaacgcttgccttaattc	362
OY	359	tggcgaatgcctgtgtttgacgtgtggaatggggaaacacagcagagggaaaacagcctgtgtccgaa	418
Db	363	tggcgaatgctgtgtatgtatgtgtgtgtgtgcacacagcagagaagaacaaactgtggccagaa	4222
OY	419	aaccacaagaacctctccatagaccacaagaataagttctcvtggtcttcagttatgatgatg	478
Db	423	gaactgtgtatattccacgaattccagtttgaattcttctgtgtgtcttcagtttgaataatg	482
OY	479	ctagttgtgaccttcctgcgcagtgataactttatcttgcagaaaacggcgcagagaaaga	538
Db	483	ctaaataatgaccttccagcccaatctgacttcatctgttaaaagaaaacctgcgcagaagcagc	542
OY	539	tctatcatctgcgtactacaaagggcaccacaaatgvgcttatgtacatttccacaatgc	598
Db	543	tacataatgtgtgcattcccaaggcgcacacattgttttatgtcctttccacaatc	602
OY	599	cagagctcgtgctcagaaaataataaattgattttgtttaagaccacatagccacatgttaagc	658
Db	603	ccagctctgtccttaaaagaatataaaacctctcatgtctctgacgtctgtgtgcacactgttaag	662
OY	659	atgcacaaaagcccgcgagccaacattttgtgtgcgcagagatagatgataagagattgt	718
Db	663	atacaaaaagccttataataacaactatgatttgttccctaatccctctcaattatata	722
OY	719	ttggcaaaaagaattctctgtatcatagaccagatttctagacaacatgttatttaacctt	778
Db	723	ttgtgtgacaataattcttaccacaacaaccttctgtatctcaattctgtctacgaagtgt	782
OY	779	gtggcaggtgattctctgtatagaattgtgtgtaatacatgattactctgtgtgtgtatca	838
Db	783	gtctccgtgtgagatgtctgaattctcttctgtgcagcaatgctcttatattataatttggattgt	842
OY	839	acaccacaataatgataacatgagccgagcaagtgatatagtctgcccacaactctgtcgtgaa	898
Db	843	acagtaagaactttaacacagatcgtctgtgattgtatcatcatcaacataatccagcagaa	902
OY	899	catctgtgcaaaaataattcttacctgtgggcccagcgatgtaattctctgtgaaattccgggcat	958
Db	903	cttctgttcaaaaataatgttccattgtgcccagcgctgtgtaagtctgtggaataattccaaagctt	962
OY	959	ttgactgtggggatgtgagaccacaataattggaanaaatgcaatccagcacaactccgttaagt	101
Db	963	atgactgtggggagcccgatccagatagatagattgcaactatgtaactgctcccaactccctact	102
OY	1019	acagagtcagagataatgacggttccctacagcaaatgtgacagaggatcgaagactgctt	107
Db	1023	acaaatgtgacagccatbaatgattglaaccaattgcagtgtggaacggtgtggcaagacactgttg	108
OY	1079	caaatctcagaagagctggaanaatgctgtctctctgaggtgagaccaaacctcatattaccacaag	113
Db	1083	ctgcgccccaagaatgttgccttctgttgcctcaaaaacctcccaactctattaccacaag	114
OY	1139	atacttccatgaaatggtctacagctgatttcaactcgggtttgtgatagtctccaccagtaatt	119
Db	1143	agattccttttatacaactcttgactttatctcgggcaatgagtgctccctccaagaagttt	120
OY	1199	acaatgaatcatcatcatctgatt	1220
Db	1203	acaatgacattgttctcatgatt	1224

	RESULT	8
AAH57420	ID	AAH57420 standard; cDNA; 1695 BP.
XX	AC	AAH57420;
XX	DT	10-SEP-2001 (first entry)
XX	DE	Human stomach cell specific cDNA sequence SEQ ID NO:260.
XX	KW	Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
XX	KV	lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
XX	KM	metabolic disease; developmental disease; cysticotic; immunomodulatory;
XX	OS	neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX	PN	Homo sapiens.
XX	PD	WO200132927-A2.
XX	PD	10-MAY-2001.
XX	PF	02-NOV-2000; 2000MO-US30396.
XX	PR	04-NOV-1999; 99US-0163508.
XX	PA	(INCY-) INCYTE GENOMICS INC.
XX	PI	Sornasse T, Seilhamer JJ, Watson GA;
XX	DR	WPI; 2001-291057/30.
PT	PT	New cell and tissue specific polynucleotides useful for diagnosis,
PT	PT	prognosis or monitoring of treatments for disorders where the gene is
XX	PS	associated with a cancer, immunopathology or neuropathology -
XX	PS	Claim 1; Page 189-190; 327bp; English.
CC	CC	AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC	CC	sequences (I). (I) can have cytosstatic, immunomodulatory and
CC	CC	neuroprotective activities, and can be used in gene therapy. (I) and
CC	CC	proteins (II) encoded by them are used in high throughput screening
CC	CC	assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC	CC	mimetics, peptides, proteins, agonists, antagonists, antibodies or
CC	CC	their fragments, immunoglobulins, inhibitors, drug compounds and
CC	CC	pharmaceutical agents. Expression of (I) in a sample indicates the
CC	CC	differentialation of embryonic stem cells into a tissue selected from
CC	CC	brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
CC	CC	tissues. (I) and (II) are used to produce an expression profile that
CC	CC	defines a metabolic or developmental process, treatment, condition,
CC	CC	disease or disorder. The gene profile can be used for diagnosis,
CC	CC	prognosis or monitoring of treatments and for investigating a
CC	CC	predisposition to a disorder where the gene is associated with a
CC	CC	cancer, immunopathology or neuropathology.
XX	XX	
SQ	SQ	Sequence 1695 BP; 480 A; 356 C; 389 G; 470 T; 0 other:
Query Match	36.9%	Score 468.4; DB 22; Length 1695;
Best Local Similarity	64.1%	Pre. NO. 1.8e-138;
Matches 706; Conservative	0;	Mismatches 386; Indels 0; Gaps 0;
OY	I19	ctgtgagaccgaagcatcattgatattagttaataatccaacataagaagtcacct 178
Dp	I16	cttggaaagcccgcaagtgactatgaacattatgcagaatgattactatggggatccc aa 175
OY	I79	gttgaggatatcgaagtgcacaatgaagaatgggtatatctttctgttaacaggatctc 238
Dp	I76	atggaagatattgaagtgtagtctgaagtggttatattcttgaagtaaatgaattcctt 235
OY	I239	gagcgctagtgcacaactaagaagcaggtctccaggcctgtgtgttactgcagcatggcc 298
Dp	I236	atgggaagaaaattcagggagatacacggccagaaacctgtgtgttttgcagcatggtt 295

Dd		500	tacatcttggttggtaattcctccagcgcaaacacatttggtttcattgacctttcttcctgaataaa	53
Oy		599	cagagcttgggcctcagaataacaaatgatatttctctcttagaccccataagccaattgtaagt	678
Dd		560	ctaacacttgcacaanaaatcaaacagccttttatgcattagctctccagtttgttaacygnagt	63
Oy		659	atgcnaaaagccccgggagcccaaattctctgttcgcgcagatatgatgatcaaggagatgct	716
Dd		620	atacacaaaagcccttgaaaaagaattccattcatcttaccatcttccttccaagcttagt	679
Oy		719	ctggcaaaaaagaattctctgatcagaccagatttctccagaaccttgatttatccctt	778
Dd		660	ctggcaagaanaaattctctctccccaacacttacttgaacttcttctgttacccgaagt	739
Oy		779	gtggccaggtgattctcttgatcagatttgtagtaataatcagttacttctcgggtgattca	838
Dd		740	gtccagggggagttctctgatctctctctctgcagcacaccttatcatcttctgttgattg	799
Oy		839	acaccaacaabatyacacatgagccgcgaagaagtgtatalgtctgccacactcttgctgaa	898
Dd		800	acaagaaaaacttaaatgtgagtcggttttgatgtgtaactatagggcataatccagagga	859
Oy		899	catctgtgcacaataatctctacacttgagccagcgagtgaaattctgtgtaactcgggcat	958
Dd		860	catctgttcacgaacttctctccactggygcacagctgtltagatctbggaattccaagct	919
Oy		959	ttgactcggggagtgagagccaanaatctgnaaaaatgacataagccaactcctgtgaagt	1018
Dd		920	tcaactcygggaagcccatcccaagaacaatgtttacatacacaacagaanaaagcctcctgat	979
Oy		1019	acagagctcagagatatatctcacttgagccagcaatgtgagcagaggttcagactgcttt	1078
Dd		980	atgattgtgtaagccatgacgtgcgtgcaggttcaggtgysgaagcgttgcaatgacatccgtg	1039
Oy		1079	caatccaagaagcgtgaaaaatgctctctctctgtaagttgaccaaactcatctcataaga	1138
Dd		1040	ctgatccccaagatgctgcacatgctctctcccaaaactccaacctcctgttccataag	1099
Oy		1139	atatctctgaattgggctccacgttgatttatctcggggttgtagtctctctaaccgtagt	1198
Dd		1100	agattcttgcctaacacacacttggacttcatctgggcaatgagtccctccaagaagttt	1159
Oy		1199	acaatgaatcatcatctgactgacagcagagagagaccaca	1240
Dd		1160	acaatgagatgatcttcattgatgtgcagaagactcaaaagatc	1201
<hr/>				
RESULT 10				
AAO68388				
ID	AAO68388	standard;	DNA; 1528 BP.	
XX	AC	AAO68388;		
XX	XX	20-FEB-1995	(first entry)	
XX	DE		Canine gastric lipase.	
XX	KW		Canine gastric lipase; CGL; dog; fat; mucoviscidosis; enzyme;	
XX	OS		bioconversion; exocrine pancreatic insufficiency; ss.	
XX	XX		Canis familiaris.	
XX	FT	Key	Location/Qualifiers	
XX	FT	CDS	1..1140	
XX	FT		/tag= a	
XX	FT	misc_RNA	1..1137	
XX	FT		/tag= b	
XX	FT		/note= "Claim 1"	
PN		MO9413816-A.		
PD		23-JUN-1994.		
XX				

16-DEC-1992:	92FR-0015201.	16-DEC-1992:	93A-C-FR0126C.
16-DEC-1992:	92FR-0015201.		
(LJOU) INST RECH JOUYETNAL.			
Benicourt C, Blanchard C, Junien J;			
WPI, 1994-217890/26.			
P-PSDB: AAR56870.			
Recombinant canine gastric lipase and nucleic acid encoding it -			
are used for improving absorption of ingested fat, treating			
mucoviscidosis etc. and in enzymatic bio-conversions			
Claim 1; Fig 8; 52pp; French.			
The sequence given below is the sequence of Figure 8, altered			
according to the amendments described on page 2 of the appended			
letter.			
CGI is used to improve absorption of ingested fat, in healthy and			
sick patients (e.g. having altered levels of gastric lipase); to			
treat conditions associated with insufficiency (or lack) of lipases,			
esp. mucoviscidosis or exocrine pancreatic insufficiency and parlic.			
where immobilised, for bioconversions, e.g. hydrolysis or			
transesterification (other mammalian gastric lipases, or derivs.,			
can be used in this application).			
Sequence 1528 BP; 394 A; 386 C; 328 G; 420 T; 0 other;			
Query Match	35.0%; Score 444.2; DB 15; Length 1528;		
Best Local Similarity	62.8%; Pred. No. 8.3e-131;		
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps			
125 acccggaagcattcagatataatgaatcattccatcaatcaatgaagcttccgttgagg	184		
26 acccggaagtgaccatgataatgaatcattccatcactcagtgaggataccagcttgagg	85		
135 aatatgaagctgcgaactgaagatggtatattccttctgttaacagagattcctcgagcc	244		
86 aatatgaagctgtgacgcgaagcaggtttatatccttggatcgcagagaattccttaaggga	145		
245 tagtcaacctgaaggaagcaggttcccaagcctgtgtgttactgtatgagatgagcctagt	304		
146 ggaataatctcgagagataatagcgcgaagaccgttgcatttttgcaaacagcttgccg	205		
305 gaggtgcagcagcagatgattcccaacccgcgcacaacaatagcctgtgcttcatctcgag	364		
206 catcgccacaacaactgattcccaacccgcgcacaacaagaagcctgttcatctccgcgcg	265		
365 atgctgtgtttgacgtgtgtgaatggggaacagcaggggaagacgcctgtgtctgaaacaca	424		
266 acgcgggttagcagctgtgtgctggggaacagcaggggcaaacctgtggcaggaagcattc	325		
425 agaaccttccatagaccacaagatgaggttctgtgcttcaagttatgatatgatatgcttagt	484		
326 tgaactactgcgccgactccgtcgatctctgtgcttcaagcttgaagagatgagtaaat	385		
485 ttgaccttcctgcagctgataaacttatttggcaaaaagcagcaggaagaaagactat	544		
386 atgaccttccgcgcacacttgaattcatcattggaagaaacgagacagagacaagctaacct	445		
545 atgtcgtatttcacagcagggcaccacatggtgcttattgcatatttccacatgccaagc	604		
446 acgttggtcattcccaaggcaccacacattgtttcatctgcctttccacacatcccaagc	505		
605 tggctcagaataataatgtatttggctttagaccccatgagcactgtttaagatgaa	664		
506 tggcgaagaacgagatacaaaccttcatgatatgactcccggttgcacacggtgaagtacacg	565		
665 aaagcccgaggacaaattcttgtgtgcgcagatatgatatgaagagattgtttggca	724		

Db 566 aaaccctgttaaaaaaacatcatgctgccttccttccttcacgttatattgaa 625
 Qy 725 aaaaagaattctgtatcagacacagattctcgaacaaactgttattacattgtgac 784
 Db 626 acaaaattcttaccacaacacactctcttgatcaattctcgcacccagagatgctcc 685
 Qy 785 aggtgattcttgatcagaattgttgatcaatcattcttccttcgggtgagatcaacca 844
 Db 686 gcgagacgttgatctctctcgcagcaacgcctctgttaccattgttgattgacacta 745
 Qy 845 acaattatgaacatgagccgagacgaagtatattgtcgcacacactcttcgtgaacatctg 904
 Db 746 tgaacttgaaacatgagctgccttgatgtgattctgtcacaataccagcaggaacacacg 805
 Qy 905 tgcacaaatcttaccatcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 964
 Db 806 ttccgaaacgtgtctccactggtccacggtgttgatgaattcgtggaagttcaagctttgact 865
 Qy 965 gggggaagtgaacacaaatctggaacaaatgcaatcagcacaactctgttaaggtacagag 1024
 Db 866 ggggaagcccgcttcagacacatgacactatcacaagacagcagcagcagcagcagcagc 925
 Qy 1025 tcagagataatgacggtccctcctacgaatgtgacagagaggtcagacgttccttcaatc 1084
 Db 926 tgcagacacatgcatgtgcacacacgagtgagacggtgacagacgttcgagcagcagcagc 985
 Qy 1085 cagaagacgtgaacaaatgctcctcctcgtgaggtgacacacacacacacacacacacac 1144
 Db 986 ctacagatgtgacactttgtcttcacaaagctccccaatctcttaccacagaagaatc 1045
 Qy 1145 ctgaatgggctcagcgtgattcattcctcgtgggttgatgctcctcactacgtatgacaatg 1204
 Db 1046 ctcttcaacacacactgacttacttctggtgcacatgagtcgctccccaagcgtttacaatg 1105
 Qy 1205 aaatcatcatcgtggt 1221
 Db 1106 aaatgttccatgacg 1122
 RESULT 11
 AAT58915
 ID AAT58915 standard; cDNA to mRNA; 1528 BP.
 AC AAT58915;
 DT 19-AUG-1997 (first entry)
 DE Dog gastric lipase coding sequence.
 KW Dog; duodenal; gastric; lipase; transgenic; plant; recombinant; extract;
 KW food; absorption; fat; pancreatic; cystic fibrosis; exocrine; dairy;
 KW hydrolysis; trans-esterification; substrate; enzyme; biofuel; ss;
 OS Canis familiaris.
 FH Key Location/Dualifiers
 FT CDS 1..1140
 FT /*tag= a
 FT /product= gastric lipase
 PN MO9633277-A2.
 PD 24-OCT-1996.
 PE 19-APR-1996; 96MO-FR00606.
 PR 20-APR-1995; 95FR-0004754.
 PA (BIOC-) BIOEM SA.
 PA (LJOU) INST RECH JOUVEINAL.
 PI Baudino S, Benicourt C, Cudrey C, Gruber V, Lenee P,
 Merot B;

XX WP1: 1996-485783/48.
 DR P-PSDB; AAM09382.
 PT Recombinant human or canine pre-duodenal lipase proth. in transgenic
 PT plants - useful for facilitating absorption of fat, as bio-catalysts
 PT and for prodn. of bio-fuel
 PS Claim 2; Fig 1; 130pp; French.
 XX This is the nucleotide sequence encoding the dog pre-duodenal (1.e.
 CC gastric) lipase enzyme. The sequence can be used to generate transgenic
 CC plants producing recombinant lipase in an enzymatically active form.
 CC Alternatively the first 12 bases of this sequence (encoding amino acids
 CC 1-4 of the protein encoded by this sequence) or bases 1-162 of this
 CC sequence (encoding amino acids 1-54) can be deleted to form the
 CC derivatives designated delta-4 or delta-54 respectively. Plants, or
 CC their extracts, expressing the lipases or the truncated derivatives, can
 CC be used: (a) as pharmaceuticals or food to facilitate absorption of fat,
 CC either in healthy subjects or in patients with inadequate levels of
 CC gastric/pancreatic lipase e.g. those with cystic fibrosis or exocrine
 CC pancreatic insufficiency, in the elderly or in patients undergoing medical
 CC treatment which alters fat adsorption; (b) for performing industrial or
 CC agricultural reactions, e.g. in processing of fats or in the dairy
 CC industry, for hydrolysis or transesterification reactions, etc., where
 CC the plant material may provide both enzyme and substrate. The transgenic
 CC plants can also be used for biofuel production.
 SO Sequence 1528 BP; 392 A; 387 C; 329 G; 420 T; 0 other;
 Query Match 35.0%; Score 444.2; DB 17; Length 1528;
 Best Local Similarity 62.8%; Pred. No. 8.3e-131;
 Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;
 Qy 125 acccgaagcattcatgatattatgtaattcacaatcacaacgaagctatccctgtgag 184
 Db 26 accctgaagtgcacatgataataaagatgatacctactggtggaacagctgag 85
 Qy 185 aatatgaagtcgacacaaatgagtggtatattcctctcgttaacagagatccctcagcc 244
 Db 86 aatatgaagttgtgacccgaacggttatattccttggtatgcaggaattcctattgga 145
 Qy 245 tagtgaacctaagaagaagcttccagcctgtgtgtgttactcagacatgctcagtgtg 304
 Db 146 ggaataattcagagataatagccgagacgtgtgcatatttgacaacaggttgcg 205
 Qy 305 gagggtgctagcactgatttccaactgcgcacaaatagcctggttcatcttcggag 364
 Db 206 calcgacacacaaactggtatctccaactgcgcacaaatagccttcatcttcggcgc 265
 Qy 365 atgctggttttagctgtgagtgaggaacagaggggaacgcccgtgtcgaacacaa 424
 Db 266 acgcccgggtgaagcgtgtgtgtgtggaacagcagggggaacacccctggccaggaagtc 325
 Qy 425 agacacttccatcagacaaagatgagttctggttctcaagtatgatatgagatgctagtc 484
 Db 326 tgtactactgcgcgcactgcgcgtcgtgattctgtggttcaagttgaagatggttaat 385
 Qy 485 ttgaccttcctgcagatgataaacttatttgcagaaaacggccaggaagaatcttt 544
 Db 386 atgaccttccgcacacatgacttcatcttgaagaaaacgggaagacagaactacact 445
 Qy 545 atgtcgtcattcacaaggaacacacatgggttattgtcatcttccacacatgcagagc 604
 Db 446 acgttggcattccagggcagcacaacattgtttatcgtcctttccacaatcccaagc 505
 Qy 605 tggctcagaanaatcaaatgtatttgccttagcaccatagccacgtcttaagatgaa 664
 Db 506 tggcgaacagatcaaaaacttcatgattagctcccggttgcacacgtgaagtagacacg 565
 Qy 665 aaagcccccggaccaaatttctgtgtcgcagacatgatatgataagaggtattgttgca 724

Db 566 aaacctgttaacaacatcagctcgtccctcgttcctccttcaagcttatttgaa 625
 QY 725 aaaaagaattctctatcagaccagattctcagacaactgttatttaactctgtgccc 784
 Db 626 acaaatattctaccacacactcttcttgatcaattctcgcacacagatgctccc 685
 QY 785 aggtgattcttgatcagatttgtagtaataatcattactctcgtgtggtgataacaca 844
 Db 666 gcgagacggtggtatcctctcgtcagcaacgcccgtttatattgttgatttgacata 745
 QY 845 acaatatgaatcagcagcagcaagtgtatagtctgcccacactcttctgtgaacatctg 904
 Db 746 tgaacttgaaatcagatgctcgtctgtatctgtctacataaaccagcaggaacatcgy 805
 QY 905 tgcataatattctcactgagcagcagcagcagtgtaattctgtgtgaactccggcatttgat 964
 Db 806 ttccagaacgtgtccacactgtccacagcgtgttaagtctggaagtcccaagcttttgact 865
 QY 965 gggggagtgagacccaataatctggaataatgcaatcagccacactcctgttaaggtacagag 1024
 Db 866 ggggaagccaggttcagacaatgacatcattacatcagagacatgctccctactacaacc 925
 QY 1025 tcagagatatcagcgtccctcaccagcaatgtgacagaggtcagagactggttcaaatc 1084
 Db 926 tgacagacatgcatggtgcacatcgcaggtgtgaaagcgtgtgcaacgactctgtgcccagc 985
 QY 1085 cagaagacgtgaaaaatgctctctctgaggtgacacactatctacataagaatattc 1144
 Db 986 ctccagatgtgaccttctgttcccaagctcccaactcattaccacaggaagattc 1045
 QY 1145 ctgaatggcctcagctgattctcctctgtggtttgtagtccctcaccgtatgtacatg 1204
 Db 1046 ctcccttaacatcacttgacttactctgtgccaatgagatgcccctcaagcgggttacaatg 1105
 QY 1205 aaatccatccatctgatg 1221
 Db 1106 aaattgttccatgatg 1122
 RESULT 12
 AAO68389
 ID AAO68389 standard; DNA; 1531 BP.
 AC AAO68389;
 XX
 XX 20-FEB-1995 (first entry)
 DT XX
 XX
 DE Canine gastric lipase (Met-CGL).
 XX
 KM Canine gastric lipase; CGL; dog; fat; mucoviscidosis; enzyme;
 KW bioconversion; exocrine pancreatic insufficiency; ss.
 XX
 OS Canis familiaris.
 XX
 XX Key Location/Qualifiers
 FT CDS 1..1143
 FT /*tag- a
 XX
 XX W09413816-A.
 XX
 XX 23-JUN-1994.
 PD
 XX 16-DEC-1993; 93MO-FR01260.
 PE
 XX 16-DEC-1992; 92FR-0015201.
 PR
 XX (LJOU) INST RECH JOUVEINAL.
 PA
 PI Benicourt C, Blanchard C, Junlen J;
 XX
 XX WPI; 1994-217890/26.
 DR P-PSDB; AAR56871.
 XX

PT Recombinant canine gastric lipase and nucleic acid encoding it -
 PT are used for improving absorption of ingested fat, treating
 PT mucoviscidosis etc. and in enzymatic bio-conversions
 PS Claim 2; Fig 8; 52pp; French.
 CC The sequence given below is the sequence of figure 8, altered
 CC according to the amendments described on page 2 of the appended
 CC letter.
 CC CGL is used to improve absorption of ingested fat, in healthy and
 CC sick patients (e.g. having altered levels of gastric lipase); to
 CC treat conditions associated with insufficiency (or lack) of lipases,
 CC esp. mucoviscidosis or exocrine pancreatic insufficiency and partic.
 CC where immobilised, for bioconversions, e.g. hydrolysis or
 CC transesterification (other mammalian gastric lipases, or derivs.,
 CC can be used in this application).
 XX
 S0 Sequence 1531 BP; 395 A; 386 C; 329 G; 421 T; 0 other;
 Query Match 35.0%; Score 444.2; DB 15; Length 1531;
 Best Local Similarity 62.8%; Pred. No. 8.3e-131;
 Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;
 QY 125 acccagaagcttatgatatattgtgaataatcacaatcagaagctatccctgtgagg 184
 Db 29 accctgaagtgacatgatatataagtcagatgatactaatggtggaatcccaagctgagg 88
 QY 185 aattgaagtgcgaactgaatgatatggtatccttctgttcaagaaggtatccagagcc 244
 Db 89 aattgaagttgtgacggaagaagcgttatatccttggtagatgcagaatctcttataggga 148
 QY 245 tagtgcacaactgaagaagcaggttccagccgtgtgtgttactgtcagacatgacctagt 304
 Db 149 ggaataatcagagatatagccggaagacctgttgcatttttgacaacagttgtccg 208
 QY 305 gaagtgctagaactgtgattccaactgcgcacacaataagcctgtggttactcttgagag 364
 Db 209 catagccacaactgtagtctcacaactgcgcacaacaagcctgtgcttcatccttgcgcg 268
 QY 365 atgctgttttgacgtgtgtagtgggaacagcaggggaagcagcctgtgtcgaaaaacaa 424
 Db 269 acgcggtgtagcaggtgtgtgtggtgggaacagcaggggaacactgtggcccgaggaaatc 328
 QY 425 agacactctccatagaccagaatgagtctgtggtcttcagtatgatatgagtgtgagt 484
 Db 329 tgtactatgcgccgactcgcgtcgtggttctggtcttcagcttgcagcagatggtcatt 388
 QY 485 ttgaccttccgtcagtgataacttattttgcagaaaacggcgcaagaaaagatctatt 544
 Db 389 atgaccttccgcgcacatltgacttcaatctggaagaaaacggcgcaagaaatcact 448
 QY 545 atgcgccttccagcagtgacccacacacatgagcttattgtgacatttccacacagcag 604
 Db 449 acgttgacatctccaggggacacacacatgtttcatgctccttccacacacacacac 508
 QY 605 tgcctcagaataataaatgatttgccttagacacacacacacacacacacacacacac 664
 Db 509 tggcgaacggtatcaaaacttctcatgcatgactgcgttgcacacacacacacacac 568
 QY 665 aaagcccgaggaacaaaattttgtgtgcagagatgatagtatgaagagattgttgaga 724
 Db 569 aaaccctgttaaaacaaactcgtcgtccctcgttcccttcaagcttataatttgaa 628
 QY 725 aaaaagaattctglatcagaccagattctcagacaactgttatttaactctgtgccc 784
 Db 629 acaaatattctaccacacacactcttcttgatcaattctcgcacacagagtgatgcc 688
 QY 785 aggtgattcttgatcagatttgtagtaataatcattactctgtgtggtatcaacaca 844
 Db 689 gcgagacggtgtagtctctctcgtcagcagcagcctgtttatattgttgattgacata 748
 QY 845 acaatatgaatcagcagcagcagtgatagtgtgtcccaactcctgtcgtgacatctg 904

Sequence 1308 BP; 402 A; 246 C; 264 G; 396 T; 0 other.

	Query Match	22.5%	Score 286	DB 22	Length 1308
	Best Local Similarity	67.4%	Pred. No. 1.8e-80		
	Matches 403	Conservative 0	Mismatches 195	Indels 0	Gaps 0
QY	86	gaaatgtcaatcagtacatgatcgtccaaacgaagcgtgtgacccagaagcattcatgaata	145		
Db	119	gattcatgtatgtgtatgatgcaagaagaagaacaatgcaaaccttgaaagctaatatgaata	178		
QY	146	ttagtgaatcattccaaacatcaaggctatccctgtgaggaatatagtgaatgtgcaactgaa	205		
Db	179	ttagccagatattattcttactcgtggttactcattagaaagatgatgtgttacaacaaag	238		
QY	206	atgggtatattccttctgttaacagagattccctcgaggccttaatgcaacttaagaagacg	265		
Db	239	atggttataatcccttggnaattttatagattcccatgtgaagagagatgcccagaagagatgc	298		
QY	266	gtcccaagcctgtgtgtgttactgcagcatgtgcctagtctgaagttccttaacaactgattc	325		
Db	299	ctccaaagccctgtgtgtatctgcagcatggtttatgatgcatctgtgcagtaactgtattc	358		
QY	326	ccaacctgtcccaacaacatgtggctgtccattctgtgcagatgtcgtgttttgaactgtgga	385		
Db	359	gcaacctgtcccaacaacagttgtgtcttccctctgtgcagatagtggttagagctgtgt	418		
QY	386	tggggaacaagcaggggaagaagcctgtgtctcgaaacacaagaagcactcttcataagaccaag	445		
Db	419	tggggaacaagcaggggaagaacacttgtgtccagaagaacacctttaatgttaccggaattcac	478		

QY 446 atgctgttcggcgtttccagtatatgaigaaagtgcagggttgcacctccgcagtgtaa 5050
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 479 cagaatacttggcccttcccgcttcttgatvgatavgtcaataatagaccttccagccacaatca 5384
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 506 acttaattcttcagaaaaacgggccagagaaaagactaatatgtctgcgttatcacaaaggca 5675
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 539 atttatcatatgaaaaaacctvgaaagaaagcactctctaactaogtvggcaccatccaabaagca 5986
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 566 ccacacatgtagctcttatttgcattttccacacatgcagagctgctcagaaaaatacaaatgt 6250
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 599 ccacacatagctttatatagcatcttctctcaaaccccagaacatgctcaaaaaagattaagatat 6586
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 636 atttgtcttagcaaccataagccactgtttaagcatgtcaaaaaagccccggagaccataatt 683
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 659 ttlttgacatgctgcagttcagttccagtttaataatcacaccaaaagtcctcatgaaaaaact 716

```

RESULT 15
AAA44349
ID AAA44349 standard; cDNA; 617 BP.

```

AC AAA443497

DT 21-AUG-2000 (first entry)

DE	Human secreted expressed tag SEQ ID NO:924
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100
101	101
102	102
103	103
104	104
105	105
106	106
107	107
108	108
109	109
110	110
111	111
112	112
113	113
114	114
115	115
116	116
117	117
118	118
119	119
120	120
121	121
122	122
123	123
124	124
125	125
126	126
127	127
128	128
129	129
130	130
131	131
132	132
133	133
134	134
135	135
136	136
137	137
138	138
139	139
140	140
141	141
142	142
143	143
144	144
145	145
146	146
147	147
148	148
149	149
150	150
151	151
152	152
153	153
154	154
155	155
156	156
157	157
158	158
159	159
160	160
161	161
162	162
163	163
164	164
165	165
166	166
167	167
168	168
169	169
170	170
171	

KM Human; mouse; chicken; rat; secreted expressed sequence tag; sEST;
KM expressed sequence tag; EST; prob; chemotactic; proliferative;
KM immunomodulatory; haematopoietic; chemokine; analgesic; haemostatic;
KM thrombolytic; antiinflammatory; cytostatic; antibacterial; antifungal;
KM antiviral; antidiabetic; antisthmatic; vulnerrary; antiparkinsonian;
KM antitumor; osteopthic; neuroprotective; nootropic; antipsoriatic;
KM cerebroprotective; anticonvulsant; antidepressant; gene therapy;
KM vaccine; autoimmune disorder; multiple sclerosis; allergic condition;
KM insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;
KM lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;
KM central nervous system disorder; Alzheimer's disease; stroke;
KM Parkinson's disease; Huntington's disease; coagulation disorder;
KM haemophilia; thrombosis; inflammatory disorder; Crohn's disease;
KM tumour; infection; depression; psoriasis; ss.

OS Homo sapiens

PN WO200021991-A1.

PD 20-APR-2000.

PF 15-OCT-1999; 99WO-US24206.

PR 15-OCT-1998; 98US-0104436.

PA (GEMY) GENETICS INST INC.

PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;

XX

XX

PT expressed sequence tags (SESTs), useful for treating various disorders.

XX

XX

CC sequence tags (SESTs), isolated from human, mouse, chicken and rat

CC the tissues they were isolated from. The activities include:

CC chemokinetic; analgesic; haemostatic; thrombolytic; antinflammatory;

CC antiasthmatic; vulnerary; antilulcer; osteopathic; neuroprotective;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2002, 20:27:03 ; Search time 90.11 Seconds
(without alignments)
3189,442 Million cell updates/sec

Title: US-09-578-063-46

Perfect score: 1269

Sequence: 1 atgttggaacctgtcaag.....gacggtgtgagccgcatatt 1269

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 351203 segs, 11323899 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq: *
6: /cgn2_6/ptodata/2/ina/Backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	444.2	35.0	1137	1 US-08-227-108-2	Sequence 2, Appl
2	444.2	35.0	1137	2 US-09-073-674-2	Sequence 2, Appl
3	444.2	35.0	1140	1 US-08-227-108-4	Sequence 4, Appl
4	444.2	35.0	1140	2 US-09-073-674-4	Sequence 4, Appl
5	444.2	35.0	1146	1 US-08-227-108-6	Sequence 6, Appl
6	444.2	35.0	1146	2 US-09-073-674-6	Sequence 6, Appl
7	444.2	35.0	1528	1 US-08-227-108-1	Sequence 1, Appl
8	444.2	35.0	1528	2 US-09-073-674-1	Sequence 1, Appl
9	48.2	3.8	178	1 US-08-751-782-2	Sequence 2, Appl
10	48.2	3.8	178	2 US-08-925-171-2	Sequence 2, Appl
11	46.4	3.7	43676	3 US-09-336-952-12	Sequence 12, Appl
12	36.6	2.9	3271	2 US-08-852-806-1	Sequence 1, Appl
13	36.6	2.9	3271	3 US-09-163-669-1	Sequence 1, Appl
14	36.6	2.9	3282	1 US-08-276-852-154	Sequence 154, App
15	36.6	2.9	3282	2 US-08-276-852-154	Sequence 154, App
16	36.6	2.9	3282	1 US-08-899-575-154	Sequence 154, App
17	36.6	2.9	3282	2 US-08-899-575-154	Sequence 154, App
18	36.6	2.9	3282	1 US-08-899-575-154	Sequence 154, App
19	36.6	2.9	3282	2 US-08-899-575-154	Sequence 154, App
20	36.6	2.9	3282	5 PCT-US95-08743-154	Sequence 154, App
21	36.6	2.9	3282	5 PCT-US95-08743-154	Sequence 154, App
22	36.6	2.9	13254	1 US-08-276-852-156	Sequence 156, App
23	36.6	2.9	13254	2 US-08-276-852-156	Sequence 156, App
24	36.6	2.9	13254	1 US-08-899-575-156	Sequence 156, App
25	36.6	2.9	13254	2 US-08-899-575-156	Sequence 156, App
26	36.6	2.9	13254	1 US-08-899-575-156	Sequence 156, App
27	36.6	2.9	13254	1 US-08-899-575-170	Sequence 170, App

28	36.6	2.9	13254	5 PCT-US95-08743-156	Sequence 156, App
29	36.6	2.9	13254	5 PCT-US95-08743-170	Sequence 170, App
30	33.6	2.6	8392	1 US-08-080-255-6	Sequence 6, Appl
31	33.6	2.6	8392	3 US-08-465-713-6	Sequence 6, Appl
32	33.6	2.6	8392	5 PCT-US93-05857-6	Sequence 6, Appl
33	32.6	2.6	7218	1 US-08-232-463-14	Sequence 14, Appl
34	32.6	2.5	2197	1 US-08-233-005-3	Sequence 3, Appl
35	32.6	2.5	2197	1 US-08-428-943-3	Sequence 3, Appl
36	32.6	2.5	2197	3 US-09-016-649-3	Sequence 3, Appl
37	32.6	2.5	2197	3 PCT-US95-04858-3	Sequence 3, Appl
38	32.6	2.5	8342	3 US-08-545-860D-63	Sequence 63, Appl
39	32.6	2.5	8342	5 PCT-US94-04496-63	Sequence 63, Appl
40	31.6	2.5	3343	4 US-08-965-762-19	Sequence 19, Appl
41	31.6	2.4	3212	3 US-08-673-814-1	Sequence 1, Appl
42	30.4	2.4	1529	1 US-08-336-778-1	Sequence 1, Appl
43	30.2	2.4	4982	3 US-08-699-103B-1	Sequence 1, Appl
44	30.2	2.4	10079	2 US-08-476-866-20	Sequence 20, Appl
45	30.2	2.4	3639	2 US-08-737-524B-26	Sequence 26, Appl

ALIGNMENTS

```

RESULT 1
US-08-227-108-2
; Sequence 2, Application US/08227108
; Patent No. 5807726
; GENERAL INFORMATION:
; APPLICANT: Blanchard, Claire
; APPLICANT: Benicourt, Claude
; APPLICANT: Junien, Jean-Louis
; TITLE OF INVENTION: Recombinant Dog Gastric Lipase
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.75
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/227,108
; FILING DATE: 03-APR-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fanucci, Allan A.
; REGISTRATION NUMBER: 30,256
; REFERENCE/DOCKET NUMBER: 7620-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8664/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1137 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1137
; US-08-227-108-2
;
; Query Match 35.0%; Score 444.2; DB 1; Length 1137;
; Best Local Similarity 62.8%; Pred. No. 1.5e-134;
; Matches 689; Conservative 0; Mismatches 408; Indels 0;

```


Db 266 ACGCGGCTAGACCTGTGGCTGGGGAACAGAGGGGCAACCTGGGGCCAGAGGAAATC 325
Oy 425 agacactctcatagaccagaatgagttctgggttcagtatgataagatgagtgatg 484
Db 326 TGTACTACTCGCCGACCTCGCTGAAATCTGGGGCTTTGACGCTTTCAGAGATGGCTAAAT 385
Oy 485 ttgacctctcgagtgatgataacttatttgcagaaaagggccaggaagaaatcatt 544
Db 386 ATGACCTTCCGCGCACCTTGAATCTTGAAGAAAACGGAGACAGCAAGCTACACT 445
Oy 545 atgtcgatctcacagggccaccacagggccttattgcaattccaccatgccaagc 604
Db 446 ACGTTGGGCATTTCCAGGGCAACACATTTGTTTATGCGCTTTCCACCAATCCCAAGC 505
Oy 605 tggctcagaanaatcaaatgtatatttgccttaagcaccatagccactgttaagcatgaa 664
Db 506 TGGCGAAACGGATCAAAACCTTCTATGCATTAGCTCCCGTTGCCACCGTGAAGTACACG 565
Oy 665 aaagcccccggagaccgaattttgtgtcgcgcagataatgataaggaattgttgca 724
Db 566 AAACCTGTAAACAACTGATGCTGCTCCCTTCCTTCAGCTTATATTGGAA 625
Oy 725 aaaaagaattctgcatcagaccagattctcagacaaactgttattacattgtgccc 784
Db 626 ACAAATATTTCTACCCACACACACTTCTTGAATCAATTTCTGCCACCGAGATGCTGCC 685
Oy 785 aggtatctctgataagattgttagtaataatcattcttctggtgagatcaaacaca 844
Db 686 GCGAAGACGGTGGATCTCTTGACCAACAGCCCTGTTTATCATTTGTGGATTGTACACTA 745
Oy 845 acaatatgaatagccaggaagtgatataatgctgcccacacttcttgtaagaacatctg 904
Db 746 TGAACCTTGAACATGATGATGCTTGGATGTATCTGTCAATATCAGAGAGAACATCGG 805
Oy 905 tgcgaatattctcacagggccagggcagtgatcttcctggtgaactccggagattgact 964
Db 806 TTCAGAAAGTGTCTCCACTGCTGCTCCAGGCTGTTAACTGCGGAAGTTCCAAAGCTTTGACT 865
Oy 965 ggggagtgagaccgaanaatctggaanaatgcaatcagcaactccctgtaagtagcagag 1024
Db 866 GGGGAGGCCAGTTTCAAGATGATGACTATCATCAGACGATGCTCCCTACTCTCAACCC 925
Oy 1025 tcaagagatagacggtccctacagcaatgtagcagggagtgtaagagtgcttcaatc 1084
Db 926 TGACGACATGATGATGCTGCAATGCGAGTGGAGACGGTGGCAACGACTTGTGGCGGAC 985
Oy 1085 cagaagacgtgaaatgctgctctgaggtgaggaacactatctacataagaatatctc 1144
Db 986 CTCACGATGTGACCTTTTGGCTTTCCAAAGCTCCCAATCTCTTTACACAGAGAAATTC 1045
Oy 1145 ctgaatgggtcagtgatattcactggtggttgatgagtcctcaccgtatgtaaatg 1204
Db 1046 CTCCTTAAATCAATCTTATCTGAGCTTATCTGGCCATGATGATGCCCTCAAGGGTTTACATG 1105
Oy 1205 aaatcatcatcgtatg 1221
Db 1106 AAATGTTTCCATGATG 1122

RESULT 3

US-08-227-108-4

Sequence 4, Application US/08227108
Patent No. 5807726

GENERAL INFORMATION:

APPLICANT: Blanchard, Claire

APPLICANT: Benicourt, Claude

APPLICANT: Junien, Jean-Louis

TITLE OF INVENTION: Recombinant Dog Gastric Lipase

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESS: Pennie & Edmonds
STREET: 1155 Avenue of the Americas

CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fannuci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1140 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1140
US-08-227-108-4

Query Match 35.0%; Score 444.2; DB 1; Length 1140;
Best Local Similarity 62.8%; Pred. No. 1.5e-134;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

Oy 125 acccagaagcattcatgataatagtgaaatcccaacatcaagagctatccctgtgagg 184
Db 29 ACCCTGGAAGTGCATGATATTAATCAATGATCACTACTGGGATACCCAGCTGAGG 88
Oy 185 aatatgaagtcgaactgaagaatggtatgatacttctgttaacagagatcctcgagcc 244
Db 89 AATATGAAGTTGTGACCCAGACGCTTATATCTTTGGATGACAGAAATTCCTATGGA 148
Oy 245 tagtcaacctgaagaagcaggttcagagcctgtgtgttactgcaagatgctagtg 304
Db 149 GGAATAATTCAAGAGAAATATAGCCCGGAGACCTGTTGCAATTTTGGCAACAGGTTGCTCG 208
Oy 305 gagggtgtagcaactggaattccaaactgcccacaataagcctggtctcaattctgagcag 364
Db 209 CATCGCCCAAAACGATGATCTCCAACTGCCCAACAGAGCTGCGCTTCAATCTCGGCGG 268
Oy 365 atgtcgtgttgaagctgtggaatggaacagcagggggaacgcctgtgtctcgaacaaca 424
Db 269 ACGCGGGGTAGACAGCTGTGCTGGGGAAACAGAGGGGCAACACCTCGGGCCAGAGAAATC 328
Oy 425 agacactctcatagaccagaatgagttctgggttcagtatgataagatgagtgatg 484
Db 329 TGTACTACTCGCCGACCTCGCTGAAATCTGGGGCTTTGACGCTTTCAGAGATGGCTAAAT 385
Oy 485 ttgacctctcgagtgatgataacttatttgcagaaaagggccaggaagaaatcatt 544
Db 389 ATGACCTTCCGCGCACCTTGAATCTTGAAGAAAACGGAGACAGCAAGCTACACT 448
Oy 545 atgtcgatctcacagggccaccacagggccttattgcaattccaccatgccaagc 604
Db 449 ACGTTGGGCATTTCCAGGGCAACACATTTGTTTATGCGCTTTCCACCAATCCCAAGC 508
Oy 605 tggctcagaanaatcaaatgtatatttgccttaagcaccatagccactgttaagcatgaa 664
Db 509 TGGCGAAACGGATCAAAACCTTCTATGCATTAGCTCCCGTTGCCACCGTGAAGTACACCG 568

Db 809 TTCAAGAGCTCTCACTGCTGCTCCAGGCTGTAAAGTCTGGGAAGTTCCAACTTTGACT 868
QY 965 ggggagtgagacaaaatctggaaaaatgcaatcagccaactcctgttaagtagagag 1024
Db 869 GGGGAGCCCGGTTCAGAACATGATGACATCATGACAGAGATGCTCCCTACTACAAAC 928
QY 1025 tcagagatagagcgtccctacagcaatgtgacagagagtgctcagagcttccaatc 1084
Db 929 TGACAGACATGCAATGCTCCATCCCACTGTGGAACGGTGCAACGACTGCTGCGCAC 988
QY 1085 cagagaagcgtgaaaaatgctgtctctgagtgagcaaacctcatcacaataaatatc 1144
Db 989 CTCACGATGTTGACCTTTGCTTTCCAAAGCTCCCAATCTCATTTACACAGAGATTTC 1048
QY 1145 ctgaatggctcagtgattcattcactcggggttggatgctcctcaccgtagtacaatg 1204
Db 1049 CTCCTTCAATCACTGACTTTATCTGGGCCATGAGATGCCCTCAACGGGTTTACATG 1108
QY 1205 aaatcatcatcgtatg 1221
Db 1109 AAATGTTTCATGATG 1125

RESULT 5
US-08-227-108-6
Sequence 6, Application US/08227108
Patent No. 5807726
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penlie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fauncel, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1146 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-227-108-6

Query Match 35.0%; Score 444.2; DB 1; Length 1146;
Best Local Similarity 62.8%; Pred. No. 1.5e-134;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 125 acccagaagcatcatgaattatagtagaatacatcacaatacgaagctatccctgtgag 184

Db 26 ACCCTGAAGTACCATGATATATAGTCAATGATGATGACCTTACTGCGGATATACCAAGCTAGG 85
QY 185 aataatgaagtcgaactcgaagtagatataccttctgttaacaggaattcctcagacc 244
Db 86 AATATGAAGTTGTATACCGAAGAGCGTTATATCCTTGGGATGCAAGATATCCTATGGA 145
QY 245 taatgcaactaagaagaacaggttccagcctgtgtgttaacttgcagacagagcctagt 304
Db 146 GGAATAATTCAGAAATAATAGCGCGGAGACCTGTTCATTTTTCACACAGGTTTGTCTG 205
QY 305 gaagtgtgaagaattgatttccaaactgcccacaataagcctggtgattcattcctgag 364
Db 206 CATACGCCACAATGATCTCAACTGCCCCAACACAGCTTGCTTCACTCTGCGCG 265
QY 365 atgtctgttttgaagctgtgtagtgaggagcagggggaacgcctgtgtctcgaataaca 424
Db 266 ACGCCGGGTAGAGAGTGTGGCTGGGGAACAGACAGGGCAACACTGGGCCAGAGATTC 325
QY 425 agaacactctcacaagacaagaatgagttctggccttcagttatgatagatagtgtagt 484
Db 326 TGTAATACTGCGCCGACCTCCGATTCGCAATTCGCGCTTTCAGCTTTGACGAGATGCTAAT 385
QY 485 ttgaccttctctgacgtatataacttatttgcagaagaacgggcaggaagaatcatt 544
Db 386 ATGACCTTCCCGCCACCATGACTTCATTTGAAGAAGACGGACAGACAGCTACACT 445
QY 545 atgtcgcattcacaagggcacacacacatgggcttatttgcatttccacatgcagagc 604
Db 446 ACCTTGGCCATTCCCAAGGACACACATTTGTTTCACTGCGCTTTCACCAATGCCAAG 505
QY 605 tggctcgaagaataaataatgtatttgccttagcaccataagcactgtttaagatgaa 664
Db 506 TGGCGAAACGGATCAAAACCTTATGACATGATGATCCCGTTGACACCGTGAGTACACCG 565
QY 665 aaagcccccggaccaaaatttgtgtgtccagatagatagatgaagatgtgtgtgga 724
Db 566 AAACCTGTTAAACAACACTGATGCTCGCTCTTGTCTCTTCAACTTATATTTGAA 625
QY 725 aaaaagaattctgtatcagacacagatttccagacacactgttatttaccctgtgtgc 784
Db 626 ACAAATATTTCTACCCACACACACTTCTTTGATCAATTTCTGGCCACGAGATAGCTGCC 685
QY 785 aggtgattcttgatcagattttagtaatatcatgttacttctgtgtgtgtgtgtgtgtgt 844
Db 686 GCGAGACGCTGATCTCTCTGACAGCAACGCCCTGTTTATCATTTGTGGATTTGACACTA 745
QY 845 acaatatgaatagcagcagcaagttatatagtctgcacacacactctgtggaacatctg 904
Db 746 TGAACCTGACATGATGCTGTGATGTGTCATCAATATCCAGACAGAACATATCGG 805
QY 905 tgcacaatatctcaactcgtgagccagggcagtgaaattctgtgtaactcgcgggcatltagt 964
Db 806 TTCAAGAGCTCTCACTGCTGCTCCAGGCTGTAAAGTCTGGAATTCACAGCTTGTACT 865
QY 965 ggggagtgagacaaaatctgaaaaatgcaatcagccaactcctgttaagtagagag 1024
Db 866 GGGGAGCCCGGTTCAGAACATGATGACATCATGACAGAGATCCTCCCTACTACAAAC 925
QY 1025 tcagagatagagcgtccctacagcaatgtgacagagagtgcaagagtgcttccaatc 1084
Db 926 TGACAGACATGCAATGCTCCATCCCACTGTGGAACGGTGGAACGACTTGTGCGCAC 985
QY 1085 cagagaagcgtgaaaaatgctgtctctgagtgagcaaacctcatcacaataaatatc 1144
Db 986 CTCACGATGTTGACCTTTGCTTTCCAAAGCTCCCAATCTCATTTACACAGAGATTTC 1045
QY 1145 ctgaatggctcagtgattcattcactcggggttggatgctcctcaccgtagtacaatg 1204
Db 1046 CTCCTTCAATCACTGACTTTATCTGGGCCATGAGATGCCCTCAACGGGTTTACATG 1105
QY 1205 aaatcatcatcgtatg 1221


```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/227,108
FILING DATE: 03-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fanucci, Allan A.
REGISTRATION NUMBER: 30,256
REFERENCE/DOCKET NUMBER: 7620-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
US-08-227-108-1

Query Match      35.0%; Score 444.2; DB 1; Length 1528;
Best Local Similarity 62.8%; Pred. No. 1.8e-134;
Matches 689; Conservative 0; Mismatches 408; Indels 0; Gaps 0;

QY 125 accagaagcatcatalgatatagtgaaatcattcaatcaatgaagctatccctgtgag 184
DB 26 ACCGAGTAGTACATGATATAGTATAGTACATGATGATCAGTACCTACTGGGATGCCAGCTGAG 85

QY 185 aataagaagtcgaactgaagaatggtatattccttcgttaacagagattcctcgaagcc 244
DB 86 AATAAGAAGTGTGACCCAGACGGCTTATCTTGGGATCGACAGAAATCTTTTGGGA 145

QY 245 tagtgaacccaagaagcaggttcacagcctgtgtgttactgaagcctgacctagt 304
DB 146 GGAATAATTCAGAGATATAGGCCGAGACCTGTTCATTTTGCACAGGTTTGCTG 205

QY 305 gagggtcagaactggaattcccaactgcccacaatagcctggtctcattctcgacg 364
DB 206 CATGAGCCACAACACTGATCTCCACACTGCCCCAACACAGCCTGGCCTTCATCTGCGCG 265

QY 365 atgtgtgttttgaagtgatgaggaagcagcaggggaaagcgtgtctcgaaacaca 424
DB 266 ACGCGGGGTAGACCTGTGGCTGGGGAAGAGAGGGGCAACACTGGGCGCAGAGGAATC 325

QY 425 agaacacttcacatagccaagaatggttctggtcttcagtatgatatgagtgtagt 484
DB 326 TGTACTACTCCGCCGACCTCCGTCGATCTGGGCTTTTCAGCTTTACGATGAGGCTAANT 385

QY 485 ttgaacttcctgcagtgataaacttattttgcagaaaaagggccagggaaaaagatcatt 544
DB 386 ATGACCTTCCCGCCACATGTGACTTCATCTGAAAGAAAAGGGAAGGCAAGCACTACACT 445

QY 545 atgtcggtatttcacagggcaccacatggtgcttattgcatlttccacattgcagagc 604
DB 446 ACGTTGGCCATTCACAGGCAACACCATGTGTTTCATCGCTTTCCACCATCCCAAGC 505

QY 605 tggctcagaataatcaaatgtattttgctttagcaccatagccactttagacatgaa 664
DB 506 TGGCGAAGACGATCAAAACCTTCTATGATGATGCTCCGTTGCCACCGTGAAGTACACG 565

QY 665 aaagcccccgggacaaatttgtgtcgcagatatgatatcaaaagggatgtttgga 724
DB 566 AAACCGTGTAAACAACTCATGTGCTGCTTCCTTCCTCAAGCTTATATTGGAA 625

QY 725 aaaaaaatttcgtatagaccagatttctcagacaactgttatttaactttgtgccc 784
DB 626 ACAAATATTTTACCCACACACTTTTGTGATCAATTTCTCGCCACGAGGATGCTCC 685

```

```

QY 785 agtgtattcctgatcagatttgtagtaatacatatcttctcgtggtgatccaacca 844
DB 686 GGGAGACGATGATCTCTCTGACAGACGCCCTTTATCATTTGTGATTTGACACTA 745

QY 845 acaatataacatgagccgaggaagtgatatactgcgcacacctctcgtgaacatctg 904
DB 746 TGAATTTGACATGAGTGCTTGATGATGATCTGATCTGATCATATATCCACAGACATGCG 805

QY 905 tgcataatattctacactggagccagcagcagcagcagcagcagcagcagcagcagcagc 964
DB 806 TTCACAGAGTGTGCTCCAGCTGTCAGGCTGCTGTTAACTGCGGAGATTCACAGCTTTGACT 865

QY 965 gggggagtgagacccaataatctgaaaaaatgcaatcagcagcagcagcagcagcagcag 1024
DB 866 GGGGAAGCCACTTTCAGACATGATGATCATATCAAGCATGAGCTGCTCCCTACTCAACCC 925

QY 1025 tcaagatatagcgttccctcagcaatgctgagcagcagcagcagcagcagcagcagcagc 1084
DB 926 TGACAGACATGATGATGCTGCAATGCGAGTGAGACGAGGCAAGCATTTGCTGGCGACCC 985

QY 1085 cagaagcgtgaaatgctgctcctcagcagcagcagcagcagcagcagcagcagcagcagc 1144
DB 986 CTCACGATGTTGACCTTTTGGCTTTCCAGCTCCCAATCTCATTTTACACAGAGATTC 1045

QY 1145 ctgaatgggtcagcgtgattcattcattcgtggttggatgctcctccatgataatg 1204
DB 1046 CTCCTTACATCATCTTGACCTTATCTGCGCATGATGAGGCCCTTAACCGGTTTCAATG 1105

QY 1205 aaatcattccattgatt 1221
DB 1106 AATTTGTTCCATGATG 1122

RESULT 8
US-09-073-674-1
Sequence 1, Application US/09073674
Patent No. 5998189
GENERAL INFORMATION:
APPLICANT: Blanchard, Claire
APPLICANT: Benicourt, Claude
APPLICANT: Junien, Jean-Louis
TITLE OF INVENTION: Recombinant Dog Gastric Lipase
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESS: Warner-Lambert Company
STREET: 2800 Plymouth Road
CITY: Ann Arbor
STATE: Michigan
COUNTRY: U.S.A.
ZIP: 48105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,674
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Crissey, Todd M.
REGISTRATION NUMBER: 37,807
REFERENCE/DOCKET NUMBER: 5072-DI-66-TMC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 734 622-1553
TELEFAX: 734 622-1553
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

```



```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,852
FILING DATE: 18-JUL-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Filling, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRL452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 154:
SEQUENCE CHARACTERISTICS:
LENGTH: 3282 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 15..452
US-08-276-852-154

```

```

Query Match      2.9%; Score 36.6; DB 1; Length 3282;
Best Local Similarity 51.5%; Pred. No. 0.12; Indels 0; Gaps 0;
Matches 84; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 435 catagaccaagatgagttctgggttcgtatgagatgagttgaccttc 494
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 302 CACAGCCTCAGTATGAGTTCAGAGCTCAGAGTCTGCAGACGCGTTTATATGTGC 361

QY 495 tgcagtgataaacttatttgcagaaaacgggccaagaaagatctatattgcgcta 554
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 362 GAGAGTGGGGCCATATAGTTGGGATGATCTCCCGAGCAATATATATGACGCTCG 421

QY 555 ttcaagggcaccacatgggttttgcatttcaccatg 597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 422 GGGCAAGAGAACCGCGTCATCTGAGCTCAGCTTCACCAAG 464

```

```

RESULT 15
US-08-276-852-169/c
Sequence 169, Application US/08276852
Patent No. 5652138
GENERAL INFORMATION:
APPLICANT: Burton, Dennis R
APPLICANT: Barbos, Carlos F
APPLICANT: Lerner, Richard A
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
NUMBER OF SEQUENCES: 170
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESS: Patent Counsel
STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,852
FILING DATE: 18-JUL-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Filling, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRL452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 169:
SEQUENCE CHARACTERISTICS:
LENGTH: 3282 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-276-852-169

```

```

Query Match      2.9%; Score 36.6; DB 1; Length 3282;
Best Local Similarity 51.5%; Pred. No. 0.12; Indels 0; Gaps 0;
Matches 84; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 435 catagaccaagatgagttctgggttcgtatgagatgagttgaccttc 494
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2981 CACAGCCTCAGTATGAGTTCAGAGCTCAGAGTCTGCAGACGCGTTTATATGTGC 2922

QY 495 tgcagtgataaacttatttgcagaaaacgggccaagaaagatctatattgcgcta 554
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2921 GAGAGTGGGGCCATATAGTTGGGATGATCTCCCGAGCAATATATATGACGCTCG 2862

QY 555 ttcaagggcaccacatgggttttgcatttcaccatg 597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2861 GGGCAAGAGAACCGCGTCATCTGAGCTCAGCTTCACCAAG 2819

```

Search completed: January 30, 2002, 22:11:14
Job time: 6251 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 30, 2002, 20:23:58 ; Search time 1536.27 Seconds
(without alignments)
8876.300 Million cell updates/sec

Title: US-09-578-063-46

Perfect score: 1269

Sequence: 1 atgttggaacctgtcaag.....gacgctgtgagccgcatatg 1269

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*
1: em_estfun:*
2: em_esthm:*
3: em_estln:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hlc:*
10: gb_estl:*
11: gb_est2:*
12: gb_hlc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rnd:*
20: em_gss_vit:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	618.2	48.7	2927	12 AK019504	AK019504 Mus muscu
2	505	39.8	954	11 BF135102	BF135102 601780142
3	497.6	39.2	1079	11 BF141567	BF141567 601788654
4	471.8	37.2	669	11 BF143546	BF143546 601790896
5	461.8	36.4	1350	12 AK010093	AK010093 Mus muscu
6	460.2	36.3	1347	12 AK010116	AK010116 Mus muscu
7	460.2	36.3	1349	12 AK009300	AK009300 Mus muscu
8	460.2	36.3	1350	12 AK010139	AK010139 Mus muscu
9	458.6	36.1	1340	12 AK010124	AK010124 Mus muscu
10	458.6	36.1	1343	12 AK009479	AK009479 Mus muscu
11	458.6	36.1	1345	12 AK009413	AK009413 Mus muscu
12	458.6	36.1	1345	12 AK009473	AK009473 Mus muscu

13	458.6	36.1	1346	12 AK009359	AK009359 Mus muscu
14	458.6	36.1	1346	12 AK009523	AK009523 Mus muscu
15	458.6	36.1	1346	12 AK009729	AK009729 Mus muscu
16	458.6	36.1	1348	12 AK010058	AK010058 Mus muscu
17	458.6	36.1	1349	12 AK009474	AK009474 Mus muscu
18	458.6	36.1	1349	12 AK009546	AK009546 Mus muscu
19	458.6	36.1	1349	12 AK009571	AK009571 Mus muscu
20	458.6	36.1	1349	12 AK009573	AK009573 Mus muscu
21	458.6	36.1	1349	12 AK010148	AK010148 Mus muscu
22	458.6	36.1	1350	12 AK009459	AK009459 Mus muscu
23	458.6	36.1	1350	12 AK009525	AK009525 Mus muscu
24	458.6	36.1	1350	12 AK009773	AK009773 Mus muscu
25	458.6	36.1	1350	12 AK010061	AK010061 Mus muscu
26	458.6	36.1	1350	12 AK010125	AK010125 Mus muscu
27	458.6	36.1	1350	12 AK010203	AK010203 Mus muscu
28	458.6	36.1	1351	12 AK009428	AK009428 Mus muscu
29	458.6	36.1	1352	12 AK009431	AK009431 Mus muscu
30	458.6	36.1	1356	12 AK010019	AK010019 Mus muscu
31	458.6	36.1	1356	12 AK010035	AK010035 Mus muscu
32	457.6	36.1	1346	12 AK009437	AK009437 Mus muscu
33	457.6	36.1	1347	12 AK010103	AK010103 Mus muscu
34	457.6	36.1	1349	12 AK010026	AK010026 Mus muscu
35	457	36.0	1346	12 AK009560	AK009560 Mus muscu
36	457	36.0	1347	12 AK010236	AK010236 Mus muscu
37	457	36.0	1348	12 AK010231	AK010231 Mus muscu
38	457	36.0	1349	12 AK009875	AK009875 Mus muscu
39	457	36.0	1350	12 AK009537	AK009537 Mus muscu
40	456	35.9	1349	12 AK010106	AK010106 Mus muscu
41	455.4	35.9	1345	12 AK009544	AK009544 Mus muscu
42	455.4	35.9	1349	12 AK009990	AK009990 Mus muscu
43	453.8	35.8	1346	12 AK010173	AK010173 Mus muscu
44	449.8	35.4	1407	12 AK009923	AK009923 Mus muscu
45	448.2	35.3	1350	12 AK010123	AK010123 Mus muscu

ALIGNMENTS

RESULT 1
AK019504
LOCUS
DEFINITION Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched library, clone:4632427C23, full insert sequence.
ACCESSION AK019504.1 GI:12859754
VERSION
KEYWORDS
SOURCE
CAP trapper.
Mus musculus (strain:C57BL/6J) 0 day neonate skin cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library
clone:4632427C23.
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2927)
REFERENCE
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Methods in enzymology. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome research. 10 (10), 1617-1630 (2000)
MEDLINE 10429374
PUBMED 11042159
REFERENCE
AUTHORS
3. (bases 1 to 2927)
Shibata,K., Itoh,M., Aizawa,K., Nagaoke,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kitunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamamoto,R., Matsunoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunura,S., Kawai,D.,

SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 954)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LLAM9242 row: c column: 12
High quality sequence stop: 695.

FEATURES
Location/Qualifiers
1..954
/organism="Mus musculus"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone_image="4008131"
/clone_lib="NCI-CGAP_Lu30"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/note="Organ: Lung; Vector: pCMV-Sport6; Site: 1; NotI; Site: 2; SalI; transgenic model MMTV-1, expression driven by MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 280 a 222 c 222 g 230 t
ORIGIN

Query Match 39.8%; Score 505; DB 11; Length 954;
Best Local Similarity 85.3%; Pred. No. 1.9e-136;
Matches 598; Conservative 0; Mismatches 100; Indels 3; Gaps 3;

QY 493 cctgcagtgataaacttatttcgcagaaaagccaggaagaatctattatgcgcg 552
DB 1 CCAGCTGTGATAAAGCTTACAGAAAAGCGGCGAAGAAAGCTATATGTCGCG 60

QY 553 tattcacaggcaccaccatgggttatttcattccaccatgcagagctgcacg 612
DB 61 TACTCACAGGCGACCCACATGGCTTATGCAATTTCCCAATGCCAGAGCTAGCTCAT 120

QY 613 aaaaacaaatgtatttcttagcacacacagcactgttaagcatgcacaaagccc 672
DB 121 AAATCAAAATGTATTTGGCTTAGCTCCTATGCACTGTATAATGCAAGAGTCTT 180

QY 673 gggaacaaatttctgtgcgcagatatatgatcaaggatgtttggcaaaaagaa 732
DB 181 GGTACCAAAATTTCTGCTGCGCAGATATGATGATTAATTTGGCAGACAAGAG 240

QY 733 ttctgtatagaccagatttccagacaactgttattacaccttggcaggatgt 792
DB 241 TTTTATACAGACTAGATTTTTCAGGACGCTTTTATTTACCTTGTGCGCAATGATT 299

QY 793 ctgtacagatttgtagtaataatcatgtactcttgggtgattcaacaacaaat 852
DB 300 CTCGACCAAAATCTGCACCAACATCTTACTCTGGAGAGTTTAAACAAACATATG 359

QY 853 aacatgagccgagcaagtgatatatgcgcacacaccttctgtgaacaatctgt 912
DB 360 AACATGAGCAGAGCAAAATGTATGTCGCCATACACCTGCTGGAACGCTGTG 419

QY 913 attctaacctggagccgaggaatcttctgtgaactccgggcttgcattgcatt 972
DB 420 ATTCTCACACGAGGAGCCGCGTAATCTTGCGGAATCTTGTGCTTGAATGGGAGT 479

QY 973 gagacaaataatctggaaaaatgcatacccaactcctgttaagtagtagtcagat 1032
DB 480 GAGACCAAAATCTGAGGAATGTATACGCCACTCTTATAGTTCAAGTTCCGAGAT 539

QY 1033 atgacggtccctcaagaatgtgacagagagtgacagagcttctcaatccagaag 1092
DB 540 ATGATGTCACCAAGCAAGTGTGACGAGAGCTGAGAGCTGAGAGCTGAGAT 598

QY 1093 gtgaagaatgtctctctgtgagtgacaaacctcatctacaataaattctgatt 1152
DB 599 GTGAAGCAAAATCTTCTGAGTAACCAACTCATCTACACAGAACATCTCGAGCTG 658

QY 1153 gctacgtagttacatctggttgatgagcttcacag 1193
DB 659 GGTATGTGGA-TTCATCTGGGGGAGTGAGTCCCTCAGCG 698

RESULT 3
LOCUS BF141567 1079 bp mRNA EST 24-OCT-2000
DEFINITION 601788654F1 NCI-CGAP_Lu30 Mus musculus cDNA clone IMAGE:4016413 5', mRNA sequence.
ACCESSION BF141567
VERSION BF141567.1 GI:10980607
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1079)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Plate: LLAM9263 row: 1 column: 14
High quality sequence stop: 663.

FEATURES
Location/Qualifiers
1..1079
/organism="Mus musculus"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone_image="4016413"
/clone_lib="NCI-CGAP_Lu30"
/tissue_type="tumor, metastatic to mammary"
/lab_host="DH10B"
/note="Organ: Lung; Vector: pCMV-Sport6; Site: 1; NotI; Site: 2; SalI; transgenic model MMTV-1, expression driven by MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 312 a 267 c 263 g 237 t
ORIGIN

Query Match 39.2%; Score 497.6; DB 11; Length 1079;
Best Local Similarity 84.1%; Pred. No. 2.8e-134;
Matches 621; Conservative 0; Mismatches 109; Indels 8; Gaps 5;

QY 256 aagaagacaggttcagagctgtgttctactcagcatgtgctagtgtgagtgctcag 315
DB 20 AAGAAGAGAGATCCAGGCGAGTGTCTACTGCAACATGCTTCTTGGGAGTGTAGC 79

QY 316 aactgagattccaaactggccaaacaatagcttggtctatcttggagagatgtgttt 375
DB 80 AACTGATTTCCACACCTGCGCACACAGCGCTGGGTTTATTTCTGCGAGATGCGTTT 139

Oy	376	gacgcgtgatgggaagcaagcagggaagacc--tgctctcgaaacaacaatcccttc	434
Db	140	GATGCTGTGGATGGAAAACAGCAGAGAAGAAACACCCTGGTCTCGAAGCACACAACCCCTCTC	199
Oy	435	catagccaaagtatgattcttggccttcagttaatatgatgatgcttaagtttaccttc	494
Db	200	CATAAGTCGAAGTAGTGTTCTGGGCTTTAGTATGATGAATAGCGTAGTTGACTTCC	259
Oy	495	tgcagtgtaaaccttatcttggcaaaaaacggcgcaagaagaagatcctaigtctgcta	554
Db	260	AGCTGTGTAATCACTTATCTTACAAAAACGGCCAGAAAAAGGCTCATATGTCGGCTA	319
Oy	555	ttcaaggcgcaaccacgaatggccttattgcatcttcccacatggccaagcttgcacaga	614
Db	320	CTCACAGGCGAACACCAATGAGGGCTTTATTGCATTTTCCAAATGCCAGAGACTAGCTCATA	379
Oy	615	aatcaaatgtatcttcttagcacccatagcaactgttaaagcatgcaaaaagccccg	674
Db	380	AATCAAAATGTATTTGGCGCTTAGCTCCATAGCCACTGTTAAATATGCAAGAAGTCTGG	439
Oy	675	gaccaaatcttgtctgcgcgaatgatgatgaagaagattgtttggcaaaaagaatt	734
Db	440	TACCAATTTCTGCTGCTGCCAGATGATGATGATCAAAGATATTGTCAGACAAGATG	499
Oy	735	tctgtataagacaagatttctcagaacaacttgtaattacaccttggccagtgattct	794
Db	500	TTTTATACCAAGACTAGATTTTTCAGGCACAGCTTTTATTACCTTTGGGCGAATGTTCT	559
Oy	795	tgatcaagatttgttagtaataatcattctctcttgggttgatcaacaacacaatatgaa	854
Db	560	CGACCAAAATCTGCACCAACATCATCTATCTCTCGGAGAGTTTAAACAAACCATATGAA	619
Oy	855	catgagccgagca--agtgatattgctcccccacactcttgcgtgagacatctgt---caa	909
Db	620	CATAGCGAGCAAGACATGTGCTCTCTGCCCTTAACCTGCTGGANACGTTTGTCGCCA	679
Oy	910	aataattcacacttgagccagg--cagtgaaattctgtg--aacccggagcatttgactgg	967
Db	680	ACCATTTCTCCACTGAGCGACGCGCGTGAATTTCTGGGCGACCTTGCTGCTTTGACTGTG	739
Oy	968	ggagtgagacaaaatat 985	
Db	740	GGCAAGCGACACACCAAGC 757	
RESULT	4		
BFI43546			
LOCUS	BFI43546	669 bp	mRNA EST 24-OCT-2000
DEFINITION	601790896F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4021870 5'		
ACCESSION	BFI43546		
VERSION	BFI43546.1	GI:10982586	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	1 (bases 1 to 669)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabs-remail.nih.gov		
	Tissue Procurement: Gilbert Smith, Ph.D.		
	CDNA Library Preparation: Life Technologies, Inc.		
	DNA Sequencing by: Inocyte Genomics, Inc.		
	Clone distribution: NCI-CGAP clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLW at:		
	http://lmage.llnl.gov		
	plate: LAM9277 row: o column: 23		
	High quality sequence stop: 659.		

FEATURES	source	Location/Qualifiers
		1..669
		/organism="Mus musculus"
		/strain="CZECH 11"
		/db_xref="taxon:10090"
		/clone="IMAGE:4021870"
		/clone_11b="NCI CGAP In30"
		/tissue_type="tumor, metastatic to mammary"
		/lab_host="DHI0B"
		/note="Organ: Lung; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; transgenic model WNT-1, expression driven by MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies.
		Investigator providing samples: Gilbert Smith, NIH"
BASE COUNT	195 a 140 c 165 g 169 t	
ORIGIN		
Query Match	37.2%; Score 471.8; DB 11; Length 669;	
Best Local Similarity	85.9%; Pred. No. 8.7e-127;	
Matches	535; Conservative 0; Mismatches 87; Indels 1; Gaps 1;	
Qy	1 atgttgsaaactctgtccaagaacagtgtatgtctcacagaatgtgaaatgtgctctg	60
Db	48 ATGTGAGAAACTCTGTCAAGAGTGTGGAGCTGTTTGCACAGATGTGAGATATGCTCTCG	107
Qy	61 attctgtgtgctgtatattctccagagaatgtgaattcaatataatccaaactaaact	120
Db	108 ATTCTGTGTAGATATTTACTCCAAAAGATGTGAAGCTGGGCAATTTCCCGAAGAACTT	167
Qy	121 gtggagccagagaagcttatgatataatagtgaatcatccaacatcaagagctatccctgt	180
Db	168 GCGGATCCAGAGACATTTATGATGATTTTGCGGAATCATCAAAACAGAGGTTATCCCACT	227
Qy	181 gaagaaatagaagtcgaactgtgaatgtgtatatacccttcctgttaagaagattccctga	240
Db	228 GAGGAGTATGAAGTGTCAACAGAAAGATGGGTACATCTTTCTGTGAACAGAAATCCTCGG	287
Qy	241 ggcctagtgcgaacctaaagaacaggttcccgagctgtgtgtgttcttcagatgtgcta	300
Db	288 GGACGACACGGTTAAAGAAAGAGAGATCCAGGCCAGTGTGTACTGCACATATGCTCTT	347
Qy	301 gttagaggtgtcagcaactgtgattccaacctgtcccaacaatagctgtgctcatctctg	360
Db	348 CTTGGGGATGCTAGCAACTGCAATTTCCAACTGCCCAACACAGCCTTGCGTTTATCTG	407
Qy	361 gcaagatgtgtgtttttgaactgttgaatgttggaacagcagcagggaaacgctgtgtctcga	420
Db	408 GCAGATGCAGG-TTTGATGTGTGATGTGGGAAACACAGAGGAAACACTGGGTCTCGGAAG	466
Qy	421 cacaagacactctccatagaaccaagaatgaatgtgtgtgttcagatataatgaatgtgtc	480
Db	467 CACAAGACCCCTCTCCATATGATCAATATAGTCTGTGGCTTTTCAGTTATGATGAAATGGCT	526
Qy	481 aggtttgaccttcccgagtgatataacttatatttcgagaaacagcgccagagaaagaatc	540
Db	527 AGGTTTGAACCTTCCACGCTGTGATTAATCTTATCTTACGAAACCGGCGAGAAAAGGTC	586
Qy	541 tatattgcggtctattccacagggcacacacatgtggtcttatatgtcaatttccacacatgtc	600
Db	587 TATTATGTGCGGCTACTCAGAGGCAACACATGAGGCTTTATTTGCAATTTTCCACATATGCA	646
Qy	601 gaactgtgtcagaaaatcaaat 623	
Db	647 GAGTAGCTCATTAACATCCAAT 669	
RESULT	5	
LOCUS	AKO10093	
DEFINITION	Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310067K20, full insert sequence.	
ACCESSION	AKO10093	
		1350 bp mRNA HTCC 05-JUL-2001

Db	372	ATGTGGCTATGATGTGTGGCTGGGGAAGAAGTCGAGGGAATACATGNGTCCGGAAAAATG	431
Qy	425	agacactctccatagaccagaatgagttctgggtcttcagtatatgatatgagcttagt	484
Db	432	TATACTAATGACCAAGACTCAGTTGAATTCGTGGCTTTTACGCTTTGATGAAAGGCTTAAT	491
Qy	485	ttagacctctcgtcagtgataaacttatttttcagaagaacggcgacgaagaagctatt	544
Db	492	ATGACCTTCCAGCCACCATTGACTTATTGTACAGAAAAGTGGACAGAGAAATACACT	551
Qy	545	atgcgcctattcacagggcaccacatctggctatttcatttccatccacatgcacagac	604
Db	552	ATGTGGTCTACCTCTCAGGSGCACACATATGGTTTTATTTCCTTTCTACCAATCCTGCTC	611
Qy	605	tgagctcagaataatcaaatgtattttgtcttagcacccatagccactgttataagcatga	664
Db	612	TGGCTAAAAAATCAAGAGGTTTATGCAATAGCTCCAGTTCCTATCTTGAAAGTATACAG	671
Qy	665	aaagccccgggaccgaattttttgtctgcgcagatatgtagtcaaggatggttggca	724
Db	672	AAATCCCTTTAAAAAGATTCTCAATTTCCTTAATGTTTCTTCCAGGAGGATATTGGTA	731
Qy	725	aaaaagattctcgatcagacacagatttctcagacaaattctcagaaatttcttaccgttg	784
Db	732	ACAAATATGTTTCATGCCCCACACACTACTTGATATCAATTTTGGTACGGAAGTGTCTAC	791
Qy	785	aggtgattcttgatcagatttgtagtaatacatgttactcttggttgattgattcaacca	844
Db	792	GGGAGCTGCTAGATCTTCTCTGACGACAGCGTTATTATCTCTTGATGGATTTGACAAGA	851
Qy	845	acaatatgaacaatgagccgagcaagtgtatagtctgcccacacctcttgctgaacactg	904
Db	852	AAACTTAATGTGAGTCGCTTGTGTGTATCTAGGCATATCCAGCAGGAAACATCTTA	911
Qy	905	tgcacaaattctacactggagccagcgagtgtaattctgtgaactcgggattgact	964
Db	912	CTCAGAGCCTTTTCCACTGGGACACCTTGCTTAATCTTGGAAGCTTCAAGCCTTAATCT	971
Qy	965	gggggaatgagacccaanaaatctggaanaatgcatcagcacaactcctgtaagtgacag	1024
Db	972	GGGGAATTCATATACAGAAATCTTACACTACATACAGAAAAGCGCTCCTACTATGATG	1031
Qy	1025	tcaagaatgtagcgtctccctacagcaatgtagcaagagtgtaagactcgtgcttcaatc	1084
Db	1032	TGTACGCCATAGACCGTGCCCAATGTCAGTGTGGAAACCGTGGCCATGACATCTGCGCTGATC	1091
Qy	1085	cagaagaacgtgaanaatgctgctctcgtgaggtgacacacatcctcacaataagaatc	1144
Db	1092	CCCAAGATGTGCGAATGCTCTTCCCAAACTCCCAACTCTTGTCATTAAGAGATTC	1151
Qy	1145	ctgaatgtagtcacgtgattcatctctgggttgtagtctccacacgtatgataatg	1204
Db	1152	TTCCTCATATCACTGAGATTTATCTGTGGCGATGATGATCGGCTCAAGAGGTTTACATG	1211
Qy	1205	aaatcatcatctgattg 1221	
Db	1212	AGATAGTTTACCATGATG 1228	
RESULT	8		
AKO10139			
LOCUS	1350 bp	mRNA	HTC
DEFINITION	Mus musculus adult male tongue cDNA, RIKEN full-length enriched		05-JUN-2001
ACCESSION	AKO10139		
VERSION	AKO10139.1	GI:12845371	
KEYWORDS	CAP trapper.		
SOURCE	Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus		

[illegible]

REFERENCE	AUTHORS	TITLE	JOURNAL	BASE COUNT	ORIGIN
4 (bases 1 to 1340)		THE RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.		393 a	
5 (bases 1 to 1340)		Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)		287 c	378 t
ADachi,J., Atzawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A., Arkawa,T., Carlinici,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K., Hirooka,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Itawa,M., Kato,K., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yasunishi,Y., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.		Direct Submission			
Submitted (10-JUL-2000)		Yoshihide Hayashizaki, the Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)			
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.					
CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGATCCAGATCCAGAGCTCTCTTTTCTTTTCTTTCN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adpctctt of sequence [5' GAGAGAGAGATTCCTCAGTTAATTAATTAATCCCCCCCCC 3']. cDNA was cleaved with XhoI and SclI. Cloning sites, 5' end: XhoI; 3' end: SclI. Host: SOLR.					
FEATURES	SOURCE	Location/Qualifiers			
1. 1340		/organism="Mus musculus"			
/strain="C57BL/6J"		/db_xref="taxon:10090"			
/db_xref="MGD:MGI:1895811"		/db_xref="MGD:MGI:191467"			
/clone="2310069D09"		/sex="male"			
/tissue="tongue"		/clone_lib="RIKEN full-length enriched mouse cDNA library"			
/dev_stage="adult"		47. 1234			
/note="putative"		/codon_start=1			
/protein_id="BA96715.1"		/db_xref="GI:12845347"			
/translation="MWLLVTVSLVSAFGAHLFGKLGPKNPPEANNAVSCMTIYWGIPSEEEVTEDEGVLIGVRIPIYGRKNSINIGKRPVAYIOLHILASATWINTLPNNSIALFIADADYVGLNISRGNVTSRKKNVYSPDSVEWMAFSPEDMAKYDLPATIDFVLTQKGOEIKIVGNKMSOGTTIGFIAFSTNPALAKIKRFYALPAVAYKTESPKKISLIRKFLIKVIRGNKMPHNYLDOFLGTEVCSRELIDLSNALFTFGGPKKLNYSRDEPVYICHHNPGCTSDLPFHAAOLAKSGKIDAYWVGSPIDMLAYNOKTQPYVVSAMTVDIIAVNGGHDIADPDVAMLPKILPNLIYHKEILLPIYHNDITRMADAPQEVYENIVTMM					
BASE COUNT	393 a	287 c	282 g	378 t	
ORIGIN					
Query Match	36.1%	Score 458.6;	DB 12;	Length 1340;	
Best Local Similarity	63.6%	Pred. No. 7.3e-123;			
Matches 699; Conservative	0;	Mismatches 399;	Indels 0;	Gaps 0;	

QY	125	accagagagattcagaatatttagaataatccaaatcaatcaagcattccctgag	184
Db	126	ACCTTGAGGAAACATGAAATGTGATGCAATATATACTTACGGGGATATCCAAATGAGG	185
QY	185	aatacgaatcgcaacttgaagatbgttatatcttctctgttaacaggaattccctcagggc	244
Db	186	AATATGAAGTGTACTGAAGATGGCTACATTCTGGGGGTATATGAAATTCCTTATGGGA	245
QY	245	taatgcacctaaagaagacaggttccaggctctggtgttactcagcagatgacctagt	304
Db	246	AGAAAAATTTGTGAATATCGGCACAGACCTGTGGCATATTTTGGACATGTTTGAATG	305
QY	305	gaagtctagcaactgatttccaacttcccaacaataagccttggtctcatcttcggag	364
Db	306	CATGAGCCAAACCTGATATTACAAACTCTCCAAACACAGCTGGCTCTCATTTACAG	365
QY	365	atgcgtgttttgcagctgtgtatgtyggacaagagggaaacgctctgtcttcgaaca	424
Db	366	ATGCTGGCTATGATGTGTGGCTGGGGAACAGTCGAGGGAATACATGATCCGGAAAAATG	425
QY	425	agaacacctccaatagaccagaatgattcttggtcttcagtatatgatatgctagt	484
Db	426	TATACTATTACACAGACTAGTTGAATTTGGCTTTCAGCTTATATAATGGCTAAT	485
QY	485	ttgacctctctgcagtgatataacttattcttcgaaaaaacggtccagaagaatcatt	544
Db	486	ATGACCTTCCAGCCACCATATGACTTATTTGTACAGAAAACTGGACAAAGATATACCT	545
QY	545	atgtcgtctatlcacagggagccaccatcaggtcttatgtcatttccaccatgcagagc	604
Db	546	ATGTGGTCACTCTCAGGGCACCATATGGTTTTATTCCTTTCTACCAATCCGTCTC	605
QY	605	tgtgtcgaataatcaaatgatttattgtctttagccaccatagccctttagcatgaa	664
Db	606	TGGCTAAAAAATCAAGAGTTTATGCAATATGCTCAAGTCTCACTGTGAAGTATACG	665
QY	665	aaagccccggagacaatttctgtctgcagataatgataatcaaggaattgtttgca	724
Db	666	AAATGCCCTTTAAAAAGATTCTACTTATTTCTATGTTTCTCTCAAGGATATTTGGTA	725
QY	725	aaaaagaattctgtatcagacagaattcttcagacaactgtttatttaactctgtggc	784
Db	726	ACAAAAATGTTCAATGCCCCAACACTACTTATGATCAATTTTGGTACGGAAAGTGTCTAC	785
QY	785	aggtgattcttgcatacagattctgtagtaataatcaatgattcttgggtgattcaaccca	844
Db	786	GGGAGCTGCTAGATCTTCTGTGACGACAGGCTTATTCATCTCTGTGGATTTCACAAGA	845
QY	845	acaatatgaacatgagccggaagaagtgtatagtctgcacacatctgtctggaacatcg	904
Db	846	AAAACTTAATGTGAGTCGGTTTGAATGTATCTAGGGCATATATCACAGAACATCTA	905
QY	905	tgcataaatcttcaacatgagccagggaggtgattctctgtgaaatccgggcaattgact	964
Db	906	CTCAAGACCTTTTCCACTGTGGCACACCTTGCTAAATCTGGGAAGCTTCAAGCTTAACT	965
QY	965	gggggagatgagaccaaaatctggaanaatgcaatcagcacaatctcgtlaagtiacagag	1024
Db	966	GGGGAATCCATTACAGAAACATGTTACTACTAACATACAGAAACGGCTCCTACTATGATG	1025
QY	1025	tcagagatagaaoggtccctacacgaatgtgtaacggaggttagaacctggttccaatc	1084
Db	1026	TGTGAGCAATGACCGTCCAAATGTCAGTGTGGACGATGCGCATACATCTCGGTGTATC	1085
QY	1085	cagaagacgttgaanaatgtctctctcgaaggtgacacaacctcatcacaataagaatatc	1144
Db	1086	CCCAAGATGTGCAATGCTGCTTCCCAAACTCCCAACTCTGTATACATTAAGGATTC	1145
QY	1145	ctgaatgggtctacagtgatttcatcttgggttttgatgcttccatccgattgtacaatg	1204
Db	1146	TTTCCCTCAATACACCTGTGACTTATGTGTGGCAATGATGATTCGGCTCAGGAGGTTTAACTG	1205

Oy	1205	aatcatccatcgtg	1221
Db	1206	AGATTGTTACCATGTATG	1222
RESULT	10		
AK009479			
LOCUS			
DEFINITION	AK009479	1343 bp mRNA	HTC 05-JUL-2001
ACCESSION		Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310022K08, full insert sequence.	
VERSION	AK009479		
KEYWORDS	AK009479.1	GI:12844303	
SOURCE		CAP trapper.	
		Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone:11b;RIKEN full-length enriched mouse cDNA library clone:2310022K08.	
ORGANISM		Mus musculus	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE		1 (bases 1 to 1343)	
AUTHORS		Carninci, P. and Hayashizaki, Y.	
TITLE		High-efficiency full-length cDNA cloning Methods in enzymology. 303, 19-44 (1999)	
JOURNAL		992792923	
PUBMED		103496636	
REFERENCE		2 (bases 1 to 1343)	
AUTHORS		Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.	
TITLE		Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
JOURNAL		Genome research. 10 (10), 1617-1630 (2000)	
PUBMED		20499374	
REFERENCE		3 (bases 1 to 1343)	
AUTHORS		Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, Y., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasai, Y., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawal, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kirio, A. and Hayashizaki, Y.	
TITLE		RIKEN integrated sequence analysis (RISA) system -384-format sequencing pipeline with 384 multipipillary sequencer	
JOURNAL		Genome research. 10 (11), 1757-1771 (2000)	
PUBMED		20530913	
REFERENCE		4 (bases 1 to 1343)	
AUTHORS		FANTOM Consortium.	
TITLE		Functional annotation of a full-length mouse cDNA collection	
JOURNAL		Nature 409, 685-690 (2001)	
PUBMED		5 (bases 1 to 1343)	
REFERENCE		Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arikawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirokawa, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Saeki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sobabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.	
TITLE		Direct Submission	
JOURNAL		Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsukuba-shi, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)	
COMMENT		Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.	

Db	911	CTCAAGACCTTTTCACATGGGACACAGCTTGGCTAAATCTGGGAGCTTCAAGCCTTA	ACT	970
Qy	965	ggggagatgaagaccacaaatctggaataatcgcaatccacatccctcgtlaagatcacag		1024
Db	971	GGGGAAAGTCATTAACAGAAACATGTTACACTCAATCAATCAAGAAACCGCTCCCTACTGTGATG		1030
Qy	1025	tcagatgatgaagctccctcctacagaagaatgtagacagggggttcagaacgtgcttcaatc		1084
Db	1031	TCTCGACCATACACCTGTGCCAATTCGAGTGTGTGAAACGGTGGCCCTACATCTCGCTGTATC		1090
Qy	1085	cagaagacgtgaaatgctgctctctgtaggtgtagaacctcatcatcacaataagatattc		1144
Db	1091	CCCAAGATGTGGCAATGCTGCTTCCCAACTCCCAACTCTGTGTACCATTAAGAGATTC		1150
Qy	1145	ctgaatggctcagctgtagttatcattcctgggttttgatgctcctcacccgtatgataa		1204
Db	1151	TTCCCTACAAATCACTCGACTGACATCTGCGGCGATGATGCGGCTCAAGAGGTTTCAATG		1210
Qy	1205	aatcatcattcattg		1221
Db	1211	AGATAGTTACCATGATG		1227
RESULT 12				
LOCUS	AK009473	1345 bp	mRNA	HTC
DEFINITION	Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310022H17, full insert sequence.			
ACCESSION	AK009473			
KEYWORDS	AK009473.1 GI:12844293			
SOURCE	CAP trapper. Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA, clone:11b;RIKEN full-length enriched mouse cDNA library clone:2310022H17.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
AUTHORS	1 (bases 1 to 1345)			
TITLE	Carninci, P. and Hayashizaki, Y.			
JOURNAL	High-efficiency full-length cDNA cloning			
MEDLINE	Methods in enzymology. 303, 19-44 (1999)			
PUBMED	99279253			
REFERENCE	10349636			
AUTHORS	2 (bases 1 to 1345)			
TITLE	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,			
JOURNAL	Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
MEDLINE	Normalization and subtraction of cap-trapper-selected cDNAs to			
PUBMED	prepare full-length cDNA libraries for rapid discovery of new genes			
REFERENCE	Genome research. 10 (10), 1617-1630 (2000)			
AUTHORS	3 (bases 1 to 1345)			
TITLE	Shibata, K., Itoh, M., Alzawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,			
JOURNAL	Kono, H., Akiyama, T., Nishi, K., Kitunai, T., Tashiro, H., Itoh, M.,			
MEDLINE	Sun, N., Ishi, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,			
PUBMED	Yamanoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,			
REFERENCE	Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,			
AUTHORS	Tone, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawal, J.,			
TITLE	Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.			
JOURNAL	RIKEN integrated sequence analysis (RISA) system-384-format			
MEDLINE	sequencing pipeline with 384 multiplexed sequencer			
PUBMED	Genome research. 10 (11), 1757-1771 (2000)			
REFERENCE	11076681			
AUTHORS	4 (bases 1 to 1345)			
TITLE	The RIKEN Genome Exploration Research Group Phase II Team and the			
JOURNAL	FANTOM Consortium.			
MEDLINE	Functional annotation of a full-length mouse cDNA collection			
PUBMED	Nature 409, 685-690 (2001)			
REFERENCE	5 (bases 1 to 1345)			
AUTHORS	Adechi, T., Alzawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,			
TITLE	Arakawa, T., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M.,			

FEATURES	source
CDs	<p>Location/Qualifiers</p> <p>1..1345</p> <p>/organism="Mus musculus"</p> <p>/strain="C57BL/6J"</p> <p>/db_xref="taxon:10090"</p> <p>/db_xref="MGD:MGI:1898391"</p> <p>/db_xref="MGD:MGI:1914967"</p> <p>/clone="2310022H17"</p> <p>/sex="male"</p> <p>/tissue_type="tongue"</p> <p>/clone_1ib="RIKEN full-length enriched mouse cDNA library"</p> <p>/dev_stage="adult"</p> <p>52..1239</p> <p>/note="putative"</p> <p>/codon_start=1</p> <p>/protein_id="BAB26312.1"</p> <p>/db_xref="gi:12844294"</p> <p>/translation="MWLLIVTSVLSAFGAGHLGKLGPKNPKANMNSOMITWGYR SESEVYTEDGYILGVYRIRPGKNSENIGRPAYIYOHGLIASATWITLPPNLSLA FILDADIDWLGNSKGTWNRKNVYISPSDEWAFSDEMATDIDLPATIDIVOKT GOEILHVHGSHQGTIGTGFIASTFNPALAKIKRPAVATVATVYTSPEFKSLIRK FLKLVNPKMFNPHNYLDOLFGTEVCSRELLDLCGNALFEICGPKKNLANSRPDV YLGNIPAGTSTODLEFHHMAOLSKGLDQAYMWSPLQMLNHYNOKTPPYDVASATPDI AYNMGSHDILADPDQVAMLPKLENNLHYKLEILFYNHLDIFWADADQEVYNEIVTWMA AED"</p>
BASE COUNT	395 a 288 c 283 g 379 t
ORIGIN	
Query Match	36.1%; Score 458.6; DB 12; Length 1345;
Best Local Similarity	63.6%; Pred. No. 7.3e-123;
Matches 698; Conservative	0; Mismatches 399; Indels 0; Gaps 0;
Oy	125 acccgagcattcatgataataatgaataatcatcaacatcaaggctacccctgtgagc 184
Db	131 ACCCGAGCAACAACTGATGTATGTCAGATGATTAACCTACGCGGAGATCCAAAGTGAGG 190
Oy	185 aatatgaatgcgaacatgaagatggatatacccttcctgtctaacaggatccctgaggcc 244
Db	191 AATATGAGTTGTACTGTAAGATGCTCATATTCGTGGGGTCTATGAAATTCCTTATGGA 250

through one round of normalization to Rot = 5.0 and subtraction to Rot = 25.0. Second strand cDNA was prepared with the primer adapter of sequence 5'-GAGGAGAGATTCGCGAGTAATTAATTAATCCCCCCCCCCC 3'. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.

FEATURES

source

Location/Qualifiers
1..1346
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/db_xref="MGI:1895805"
/db_xref="MGI:1914967"
/clone="2310015H15"
/sex="male"
/tissue="tongue"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
53..1240
/note="putative"
/codon_start=1
/protein_id="BAB26240.1"
/db_xref="gi:12844109"
/translation="MWLLIVSVLSAFGAGLGGKLPKPEANMNSQMTTWGIP
SEEEVETEDYILGVRIPIYKKNSENIGRPVAYLQHLASAKNITNLPNNSLA
FILDAGYDVLGNSRGNTWSRKNVYSPDSVEAFSEFDMAXYDLPAIDFIVK
GOEKIHYHGSGTIGFIAFSTNPALAKIKRFALAPVATVKTSEPFKIHIP
FLKVIENGKEMPHNYLDFLGEVCSRELIDLSALFIPGSPDKKNVSRFDY
YLGHNPACTSDDEPHNQALAKSKGLQYNNMSPLONNLHNQTPPYDVSAMTVPI
AVMNGHDILDPDQVAALPLRPLPLTYHKEILPYNHLDFIWMADAPQEVNIVTM
AED"

CDS

BASE COUNT 396 a 289 c 282 g 379 t
ORIGIN
125 acccgaagcattcatgaattatgtaatacatcaacaatacgaagctatcgtgaag 184
132 accctgaagcgaacatgtaattgtagcatgatacttaactgaggatgacgaag 191
185 aatatgaagtcgaactgaagaatggtatatactcttctgtaacaggaattccgagcc 244
192 aatatagaattgtaactgaagatgctacatctggcggtctatagaattccttgaaga 251
245 tagtgaactgaagaagaaggtccagagcctggtgtgttactgaagcagctagtg 304
252 agaaaaattctgagatatacggcaagacgctggtgcatatttgacagcatggttgattg 311
305 gaggtgctgaactggaattcccaactgccaacaatagctggtgcttcatcttcgagc 364
312 catcagcccaaaaactggaattacaaatctgcccacaaacagccttgaccttcatctgaacg 371
365 atgctggttttgacgttgatggggaacagcaggggaacgcctggttcgaaaaaca 424
372 atgctggttatgattgctggcggaacagcaggggaatgacatgctcccgaaaaatg 431
425 agacactcccatagacaagaatgagttgctggtcttcaagtatgataagatgctagtg 484
432 ttactatttcacacagctcagttggtgcttgccttgcattgataaagtgcttaaat 491
485 ttgaccttcctcagtgataaacttatttgcagaaaaacgagcaggaagaaactcatt 544
492 atgaccttcac 551
545 atgctgctattcagagggcagcagcagcagcagcagcagcagcagcagcagcagc 604
552 atgctgctattcagagggcagcagcagcagcagcagcagcagcagcagcagcagc 611
605 tggctcagaataatcagaatgatttgcctttagcacccatagccactgttaagctgaa 664

Query Match 36.1%; Score 458.6; DB 12; Length 1346;
Best Local Similarity 63.6%; Pred. No. 7.3e-123;
Matches 698; Conservative 0; Mismatches 399; Indels 0; Gaps 0;

125 acccgaagcattcatgaattatgtaatacatcaacaatacgaagctatcgtgaag 184
132 accctgaagcgaacatgtaattgtagcatgatacttaactgaggatgacgaag 191
185 aatatgaagtcgaactgaagaatggtatatactcttctgtaacaggaattccgagcc 244
192 aatatagaattgtaactgaagatgctacatctggcggtctatagaattccttgaaga 251
245 tagtgaactgaagaagaaggtccagagcctggtgtgttactgaagcagctagtg 304
252 agaaaaattctgagatatacggcaagacgctggtgcatatttgacagcatggttgattg 311
305 gaggtgctgaactggaattcccaactgccaacaatagctggtgcttcatcttcgagc 364
312 catcagcccaaaaactggaattacaaatctgcccacaaacagccttgaccttcatctgaacg 371
365 atgctggttttgacgttgatggggaacagcaggggaacgcctggttcgaaaaaca 424
372 atgctggttatgattgctggcggaacagcaggggaatgacatgctcccgaaaaatg 431
425 agacactcccatagacaagaatgagttgctggtcttcaagtatgataagatgctagtg 484
432 ttactatttcacacagctcagttggtgcttgccttgcattgataaagtgcttaaat 491
485 ttgaccttcctcagtgataaacttatttgcagaaaaacgagcaggaagaaactcatt 544
492 atgaccttcac 551
545 atgctgctattcagagggcagcagcagcagcagcagcagcagcagcagcagcagc 604
552 atgctgctattcagagggcagcagcagcagcagcagcagcagcagcagcagcagc 611
605 tggctcagaataatcagaatgatttgcctttagcacccatagccactgttaagctgaa 664

Db 612 TGGCTAAAAAATCAAGAGTTTATGCAATAGCTCCAGTGGCTACTGTAAGATACAG 671
Oy 665 aaagcccgaggaaccaatttctgtcgtccagatagatcatcaaggatgttgcga 724
Db 672 AAAGTCCCTTTAAAAAGATTCTACTTATTCCTCAAGATTCTCTCAAGGTGATTTGGTA 731
Oy 725 aaaaagaatttctgtatcagaacagaattcctcagaacattgttatccttctgccc 784
Db 732 ACAAAATGTTTCCATGCCCCAACACTAGATCAATTTCTGGTGGGAGGTGCTCAG 791
Oy 765 aggtatcttctgatacagaattgtagtaataatcatgattacttctggtgattcaacca 844
Db 762 GGGAGCTGCTAGATTTCTCTGACCAACGCTTTATCTCTCTGTTGATTTGCAAGA 851
Oy 845 acaatataacatgagccgagcaagtgatatactgctcccaactcttctgtaacatctg 904
Db 852 AAAACTTAATGATGAGTGTGCTTTGATGTGATCTAGGAGCATATGACAGAGACATCTA 911
Oy 905 tgcagaataattctacactgtagcagcaggaagtaattctgtgtaactccgagcattgact 964
Db 912 CTCAGACCTTTTCCATGCGGACAGCTTGTGTAATCTGGGAGCTTCAAGCCTTAACT 971
Oy 965 ggggagagagagacccaaatctggaataatgcaatcagcccaactctctgaggtacag 1024
Db 972 GGGGAGAGTCCATTCACAGAACTGTTACACTCAATCAACAAACGCTCTCTACTTGTATG 1031
Oy 1025 tcaagatataacgcttccctcagaagaatgtagcaggaagcagcagcagcagcagc 1084
Db 1032 TGTCCGCAATGACCTGCGCAATTCGATGCTGAGAACGCTGCAATCTCTGCTGATC 1091
Oy 1085 cagaagagctgaataatgctctctctctgtagcaggaacccatcattacataagaatctc 1144
Db 1092 CCCAAGATGTGCGCATGCTGCTTCCCAAACTCCCAACCTCTGTGATGAGGATTC 1151
Oy 1145 ctgaatgggctcagctgatttcaatcgtggtttgagatgctcctcattgataatg 1204
Db 1152 TTCCTTCAATCCTCTGAGCTTCACTGAGGAGTGGGAGTGGCTTCAAGAGGTTACAAATG 1211
Oy 1205 aaatcattcattgattg 1221
Db 1212 AGATGCTTACCATGATG 1228

RESULT 14
AK009523
LOCUS AK009523 1346 bp mRNA HTC 05-JUL-2001
DEFINITION Mus musculus adult male tongue cDNA, RIKEN full-length enriched library, clone:2310028109, full insert sequence.
ACCESSION AK009523
VERSION AK009523.1 GI:12844369
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain: C57BL/6J) adult male tongue cDNA to mRNA, clone:2310028109.
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1346)
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning methods in enzymology. 303, 19-44 (1999)
JOURNAL MEDLINE 99279253
PUBMED 10349636
REFERENCE 2 (bases 1 to 1346)
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL MEDLINE 20499374
PUBMED 11042159
REFERENCE 3 (bases 1 to 1346)
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,

